

USAGE-BASED APPROACHES TO DUTCH

Usage-Based Approaches to Dutch

Lexicon, grammar, discourse

Arie Verhagen & Jeroen van de Weijer (eds.)

LOT
Utrecht
2003

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6000
e-mail: lot@let.uu.nl
<http://www.lot.let.uu.nl/>

Cover illustration: Windmills at Kinderdijk

ISBN 90-76864-47-0
NUR 632

© copyright 2003 by the individual authors.

Usage-based approaches to Dutch

Lexicon, grammar, discourse

Arie Verhagen and Jeroen van de Weijer (eds.)

Contents

Introduction	1
1. Ariane van Santen (Leiden University) How feminine is a linguist? On the meaning of non-feminine names of professions	7
2. Arie Verhagen (Leiden University) The Dutch <i>way</i>	27
3. Robert S. Kirsner (University of California at Los Angeles) On the Interaction of the Dutch Pragmatic Particles <i>hoor</i> and <i>hè</i> with the Imperative and Infinitivus Pro Imperativo	59
4. Judith Loewenthal (Leiden University) Meaning and use of causeless causative constructions with <i>laten</i> in Dutch	97
5. Thomas F. Shannon (University of California at Berkeley) Drift in Dutch: Fleshing out the factors of change	131
6. Liesbeth Degand (University of Louvain, Louvain-la-Neuve) & Henk Pander Maat (Utrecht University) A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale	175
Index	201

Introduction

Arie Verhagen (Leiden University) & Jeroen van de Weijer (Leiden University)

On March 22 and 23, 2000, Leiden University hosted an international symposium under the title ‘Language, Culture, and Cognition’. The organizers (Vincent van Heuven, Jeroen van de Weijer, and José Birker) had decided that the presence in Leiden of a number of interesting linguists at the time provided an opportunity for a workshop that should not be passed up. They succeeded in persuading the participants to speak at the symposium, and Leiden University to sponsor it. At the end of the second day, the organizers and the participants agreed that despite the diversity of the topics addressed, there had definitely been more than one common underlying thread. One of these was the central position assigned to meaning and function in explaining linguistic phenomena, another the large role played by the analysis of actual utterances to lay an empirical foundation for the hypotheses being investigated, and a third one the ambition to account for characteristics of linguistic systems and for linguistic usage simultaneously, using the same conceptual tools. It was this sense of commonality that made us conclude that we should definitely pursue the somewhat vague plan for a volume that we already had before the symposium.

The diversity just mentioned concerned the levels of linguistic description, which ranged from that of bound morphemes (the smallest meaning bearing units) to texts (the most composite sort of structural and functional linguistic units). But many contributions to the symposium testified to the idea that the application of some form of usage-based approach actually reveals cross-level similarities. Slightly twisting an English proverb, one of these similarities that is arguably most important may be summarized as: ‘The sting is in the detail’. As scientists, linguists aim for broad, possibly grandiose generalizations, as scientists should. By the same token, however, they should also aim for accounting for a

maximum range of phenomena (with a minimum of conceptual distinctions). Especially since the development of computerized corpora has made the level of detail in patterns of linguistic usage much more amenable to investigation, paying attention to these details more often than not leads to the insight that something important was overlooked in the relatively abstract characterization that had been set up originally.

The original plan for this volume, as conceived immediately after the symposium, was to have a collection of papers demonstrating these points at different levels of description and in different languages. However, it became apparent after some time that the second part of this goal would be hard to achieve. We then decided to concentrate on a single language, viz. Dutch, rather than have a collection that would for the most part present material from one language, with a small but still mixed bag of contributions about a few other phenomena in a few other languages as a kind of appendix. The result is the present volume of six papers. It comprises a number of studies on relationships between the general and the specific in one language for a variety of domains, showing both individually as well as collectively that the specifics, although usually compatible with the general, cannot be taken for granted.

The part of our original plan that involved covering different levels of description has been maintained, and determines the organization of the volume. Starting with the lexicon, Ariane van Santen ('How feminine is a linguist?') takes her point of departure in the observation of an aspect in which Dutch differs both structurally and in terms of usage from English as well as German. This concerns the availability of several morphemes for feminine names of professions in the language and the way these are used, in relation to the way non-feminine names (among which the morphologically simplex ones) are used. The crucial issue is the question of the meaning of the *non*-feminine names. Van Santen discusses a number of more or less traditional semantic distinctions that may help to shed light on this issue; each has its merits, but in the end a full usage-based perspective is unavoidable if one seeks to account for the whole range of phenomena encountered. Ultimately, Van Santen concludes: "it is not the[] meaning [of words] that dictates the possibilities to use them, but the other way around: actual use is decisive for the meaning of non-feminine names of professions."

The study by Arie Verhagen ('The Dutch *way*') involves both specific lexical items as well as a specific grammatical construction.

Verhagen starts with a discussion of the Dutch analog of the well known English *way*-construction, showing both parallels and differences between Dutch and English. Besides obvious similarities, there are lexical, structural, and usage differences. The Dutch default lexical verb (*banen*) in the construction is also specific to it; it contains a reflexive pronoun and not a possessive one like English; the conditions for the proper use of the constructions do not fully overlap either. The latter two points turn out to be related to another difference between English and Dutch: the Dutch construction is a member of a small ‘family’ of constructions that does not seem to be a part of the grammar of English. In his conclusion, Verhagen discusses possible consequences of his analysis for very general ideas about the organization of the grammar of a language.

Moving ‘up’ further from words into grammar, Robert Kirsner (‘On the Interaction of the Dutch Pragmatic Particles *hoor* and *hè* with the Imperative and Infinitivus Pro Imperativo’) addresses a similar issue: the interaction between certain lexical items – in this case, two pragmatic particles – and certain grammatical constructions – two semantically distinct forms, each highly schematic, for issuing directive speech acts. On the basis of previous analyses of both the particles and the different imperative patterns, Kirsner derives a number of predictions about their interaction, i.e. going from a general to a specific level of description. He then tests these predictions against experimental as well as corpus data. Interestingly, the results not only corroborate the basic predictions, but also yield a number of unexpected results, which, on further scrutiny, point to the insight that *collocations* have different degrees of idiomaticity, and that some of them “lead a life of their own and have – almost like lexical items – emergent properties which are not entirely predictable from those of their components”. Kirsner also concludes by pointing out how the study of such details bears directly on ‘big’ general questions of the structure and function of language.

At a still higher level of abstractness, Judith Loewenthal (‘Meaning and use of causeless causative constructions with *laten* in Dutch’) considers the question of how the meaning of a general construction (the main Dutch causative construction) interacts with the way participants in the event are realized syntactically – or rather, *not* realized, as she specifically focuses on instances of use of the causative construction in which the causee-role is not filled (as in *Hij laat zijn huis overschilderen*, lit.: He lets his house repaint, i.e.: ‘He is having his house repainted.’).

Starting from previous studies, she first of all explores the ways in which the semantics of participant expressions and complement verbs interacts with the meaning of the causative construction. She thus shows that such previous studies actually give rise to conflicting predictions about the overall meaning of a causative clause from which an explicit causee is absent. Using actual usage data, she is able to resolve the conflict – but only by distinguishing systematically between different *subtypes* of such clauses. For example, a permissive reading of a causeless causative occurs only with a reflexive ‘affectee’ (object of the complement verb), while a nominal (especially inanimate) affectee appears to correlate with a coercive reading. Each of these and other specific subpatterns turns out to exhibit other semantic and pragmatic regularities as well.

The importance of regularities at the level of specific *combinations* of linguistic units also comes out clearly in the study by Thomas Shannon (‘Drift in Dutch: Fleshing out the factors of change’), on the changes in the relative ordering of subjects and pronominal objects in the history of Dutch since the 16th century – a ‘purely’ syntactic issue, if ever there was one. Shannon’s aim is to provide an explanation for what appears to be a rather radical change in the syntax of Dutch. Starting from a situation in which pronominal objects *as a rule* precede nominal subjects, Dutch has moved to a situation for which the reverse is true: in the modern language pronominal objects *as a rule follow* nominal subjects. As is well known – and Shannon strengthens the point with a wealth of evidence –, the change in no way occurred ‘overnight’, as if some switch was magically flicked. Although the ‘initial’ and the ‘final’ stage may give the appearance as if a rule has been replaced by its converse, close inspection of actual data from texts shows that the conditions of the linear ordering involve specific properties of both the pronominal objects (such as type: personal, reflexive, demonstrative; and case: dative, accusative) and the nominal subjects (such as definiteness, length, animacy). Actual utterances by speakers always involve combinations of such features, and it is a speaker’s estimate of the communicative success of her entire utterance that determines which order is used, thus producing, and over time changing, regularities in the grammar.

In his conclusion, Shannon draws attention to the fact that the type of usage-based explanation for grammatical regularities that he envisages, may also shed light on differences between parts of the grammar of closely related languages, such as Dutch and German. In other studies in this volume, a comparative dimension is sometimes also present, more or

less explicitly (e.g. Van Santen on usage of feminine and non-feminine nouns in Dutch as opposed to German and English, Verhagen on the Dutch *weg* construction as compared to the English *way* construction, Kirsner on pragmatic particles and imperative structures of Dutch not found in English). In each case, the comparative perspective confirms the relevance of the attention to details. Apparent similarities at a general level consistently seem to hide important differences, which only become obvious when one turns to the specifics.

This comparative dimension is a central concern in the last study of this volume by Liesbeth Degand and Henk Pander Maat ('A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale'). At the same time, this is also the study that explicitly concentrates fully on the analysis of structure and function at the most comprehensive level of language use, that of coherence relations in texts. In large corpora of French and Dutch newspaper texts, they chart the discursive contexts in which distinct 'backward' causal connectives naturally occur. On this basis, they are able to uncover "subtle meaning differences within a language but also cross-linguistically", which can nevertheless be conceptually connected: Degand and Pander Maat propose a general conceptual scale of 'speaker involvement', from which particular connectives in a particular language pick particular values. Thus they contribute to new possibilities for conceptualizing relations between general and specific dimensions of language, in this case especially in the domain of semantics and pragmatics.

The studies collected here vary not only in the specific topics they address, but to some extent also in the theoretical concepts and assumptions employed to analyze the phenomena. Communication between linguists using distinct theoretical instruments may lead to convergence on the theoretical level as well, especially if they share a fundamental view of language; we would be pleased if the present volume helped to advance such a convergence. On the other hand, the diversity of linguistic phenomena at the level of specifics is actually quite astonishing, so one should perhaps expect that a certain diversity of analytic tools is simply necessary. In that spirit, we would like to present the different perspectives offered here as contributing, individually and collectively, to the fulfillment of the common ambition to understand languages – in the present case especially the Dutch language – as they are actually used.

Acknowledgements

We would like to thank the Faculty of Arts & Humanities of Leiden University for sponsoring the March 2000 symposium on ‘Language, Culture, and Cognition’, and for thus contributing to the production of this volume. Thanks are also due to the Netherlands Institute for Advanced Study (NIAS) in Wassenaar; the editorial work was completed in the academic year 2002/2003, when we were both Fellows-in-Residence at NIAS.

How feminine is a linguist? On the meaning of non-feminine names of professions

Ariane van Santen (Leiden University)

1. Introduction

Dutch, unlike English, has quite a number of suffixes that can be used to form female personal nouns.¹ With *-e* new names can be formed on the basis of non-feminine names, as is shown in (1):

(1)	pedagoge	pedagoog	‘(female) pedagogue’
	violiste	violist	‘(female) violin-player’
	studente	student	‘(female) student’
	fotografe	fotograaf	‘(female) photographer’

Another very productive suffix is *-ster*, which, like *-e*, can be attached to non-feminine names as in (2), but can also form new names by a process of substitution of *-er*, as in (3):

(2)	tuinierster	tuinier	‘(female) gardener’
	luisteraarster	luisteraar	‘(female) listener’
(3)	zwemster	zwemmer	‘(female) swimmer’
	deeltijdster	deeltijder	‘(female) parttimer’

Although there are cases in which the speaker may be in doubt about which suffix to attach (as for instance, *brigadiere* or *brigadierster* ‘female police-officer’), in general the distribution of *-e* and *-ster* is clear: *-e* is attached to non-native names with stress on the last syllable, *-ster* to non-feminine names with *-ier* or *-aar*, or names with *-er* which, in turn, are derived from verbs (like *zwemmen* ‘to swim’) or nouns (like *deeltijd* ‘parttime’).

¹ Cf. de Caluwe & van Santen (2001) for a detailed study of the possibilities for the formation of female personal names in Dutch.

In addition to these two productive suffixes, there are many feminine names with suffixes that are no longer productive, such as the native suffixes *-es* and *-in*, as in (4):

- | | | | |
|-----|---------|--------|--------------------|
| (4) | cheffin | chef | ‘(female) chief’ |
| | lerares | leraar | ‘(female) teacher’ |

There are also non-native suffixes, which appear in foreign words, as in (5):

- | | | | |
|-----|---------------|-------------|----------------------|
| (5) | adviseuse | adviseur | ‘(female) advisor’ |
| | redactrice | redacteur | ‘(female) editor’ |
| | organisatrice | organisator | ‘(female) organizer’ |
| | historica | historicus | ‘(female) historian’ |

Besides these derivations, there are compounds, for instance with *-vrouw*, *-dame*, *-zuster* or *-meisje*, such as:

- | | | | |
|-----|------------------|------------------|------------------------|
| (6) | zakenvrouw | zakenman | ‘(female) businessman’ |
| | bardame | barman | ‘(female) barman’ |
| | ziekenzuster | ziekenbroeder | ‘(female) nurse’ |
| | kostschoolmeisje | kostschooljongen | ‘(female) boarder’ |

All in all, this means that in talking about women or female professionals in particular, in Dutch we *can* use ‘their own’ term. Frequently, however, the non-feminine name is used. In this respect, the situation is different from German, where for women more often a feminine name, in *-in*, is used: *Direktorin* (‘female director’), *Schriftstellerin* (‘female writer’). How is this possible, and why do the Dutch do it? These are the questions I would like to address in this paper.

The clue to the answer lies, without doubt, in the meaning of the non-feminine names. In Holland, there is an ongoing debate about the use of names for professions. One option is to *differentiate*, and use gender-specific names consistently. The other option is to *neutralize* gender distinctions, i.e. use one gender-neutral name for a man, for a woman, and in all those cases in which gender, as irrelevant, is not expressed.² I

² De Caluwe & van Santen (2001) discuss at length the linguistic and social

do not wish to discuss all advantages and disadvantages, but concentrate especially on the meaning and use of non-feminine names. What exactly is their meaning? There is no problem with the meaning of the feminine names – they are ‘female’ – but things are less clear with the non-feminine ones. Are they male, or indeed just non-female, i.e. gender-neutral? In other words: how feminine are the non-female names? Consider the sentence in (7):

- (7) Drie linguïsten hadden hun man meegenomen naar het congres.
‘Three linguists had brought their husbands to the conference.’

Of course one can *say* (7), but in contrast to our response to sentence (8), we may be somewhat surprised:

- (8) Drie linguïsten hadden hun vrouw meegenomen naar het congres.
‘Three linguists had brought their wives to the conference.’

In Dutch, the interpretation of (8) goes unnoticed, or automatically, while in (7) we are, at first, momentarily deceived, and then there is the reaction: “Oh yes, of course, we are dealing with female linguists here.” For this reason, certain feminist linguists call the gender-neutral quality of terms such as *linguïst* a myth (van Alphen 1983:310) or they say that gender-neutral personal names, or names for professions do not exist (Huisman 1985:70). Are they right or are they not? In other words, what do non-feminine names for professions signify? On present showing it is, unfortunately, already 1-0: the non-feminine names are more interesting than their feminine counterparts.

2. Two types of opposition

Let me first distinguish two types of opposition between feminine and non-feminine names. As (9) shows, *linguïste* is exclusively female: we can only use it with women, never with men:

background and implications of the two options.

- (9) *Hij is linguïste.
‘He is a linguist[+fem].’

The gender-neutral function is also not an option. Sentence (10) can only be read as a demand for a female linguist:

- (10) Er is een vacature voor een linguïste.
‘There is a vacancy for a linguist [+fem].’

On the other hand, *linguïst* is not exclusively male. We can use it when we are talking about a woman (11) or to refer to people, irrespective of gender (12) and we can also use it when we are talking about a man (13) or men (14):

- (11) Lisa Cheng is kort geleden benoemd als linguïst bij de Opleiding Algemene Taalwetenschap.
‘Lisa Cheng has recently been appointed as a linguist at the Department of General Linguistics.’
- (12) Er zijn veertien linguïsten uitgenodigd om op dit congres een lezing te houden.
‘Fourteen linguists have been invited to give a lecture at this conference.’
- (13) Is het toeval of niet dat er precies evenveel linguïsten als linguïstes zijn?
‘Is it a coincidence or not that there are precisely as many linguists as there are linguists[+fem]?’
- (14) Aanleiding voor dit congres is de oratie van de Nederlandse linguïst Arie Verhagen.
‘The occasion for this conference is the inauguration of the Dutch linguist Arie Verhagen.’

These possible usages are usually accounted for by means of the assumption that *linguïste* contains the component [female], which is absent from the corresponding name *linguïst*. It is therefore marked, not only morphologically (*linguïst* + *-e*), but also semantically, the non-feminine name being the unmarked member of this opposition. This type of opposition is called a privative one. Of vital importance is the fact that *linguïst* lacks the feature [male], as is shown by sentences such as in (11). It is

very well possible for us to refer to a woman by saying that she is a *linguïst*.

This is not the case with *verpleger* ('male nurse') and *verpleegster* ('nurse'). In Dutch, the sentence in (15) is ill-formed:

- (15) *Zij is verpleger.
'She is a male nurse.'

It should be:

- (16) Zij is verpleegster.
'She is a nurse.'

Apparently *verpleger* has the feature [male]. *Verpleger* and *verpleegster* differ from each other in a single semantic feature. While *verpleger* is [male], *verpleegster* is [female]; they form a so-called equipollent opposition.

The two types of contrast may be summarized as in (17):

- (17)
- | | |
|---|---|
| <pre> equipollent / \ verpleger verpleegster male female </pre> | <pre> privative / \ linguïst linguïste - female </pre> |
|---|---|

We can approach the two kinds of opposition in yet another way. As pointed out above, the two words involved in an equipollent opposition differ from each other in one respect while being semantically identical in all others. But sometimes this very difference is not important. You may, for instance, want to know how many children someone has without referring specifically to boys or girls. So your question would be: "How many children do you have?" rather than "How many boys and/or girls do you have?" Which word do the Dutch use in case they prefer not to specify the gender of a person?

As far as the names *verpleger* and *verpleegster* are concerned, there is always the option of resorting to a separate word – *verpleegkundige* – in case we prefer not to distinguish by gender. We could similarly use a separate word to replace the pair *linguïst-linguïste*, namely *taalkundige*, but this is not really necessary: because *linguïst* lacks a

gender-based component, it can be used in cases where the male-female opposition is removed:

- (18) *verpleegkundige* *linguïst¹/taalkundige*
 / / /
 verpleger *verpleegster* *linguïst²* *linguïste*

In other words, *linguïst* is the more specific term and the more general term all in one. Or, in the terminology of Lyons (1969:454): *linguïst²* and *linguïste* are co-hyponyms of *linguïst¹ = taalkundige*. This double function of *linguïst* is characteristic for members of unmarked categories. Before moving to the specific issue of the meaning of non-feminine names, it is therefore useful to have a look at the meaning of unmarked categories in general.

3. The meaning of unmarked categories

Jakobson defines the meaning of marked and unmarked categories as follows:

- (19) “The general meaning of a marked category states the presence of a certain (whether positive or negative) property A: the general meaning of the corresponding unmarked category states nothing about the presence of A and is used chiefly but not exclusively to indicate the absence of A”. (1957:5)

The first example of a marked category that I turn to concerns the diminutives as opposed to their base noun, as in *kamertje* (‘small room’) versus *kamer* (‘room’). In addition to the various connotations that they also have, diminutives denote ‘smallness’, i.e. they contain the sense-component [small], which is lacking in the positive. We can use *kamer* when we do not wish to point out that the room concerned is in fact a small room, but when used in opposition to *kamertje*, it could just as well refer to a room of a more than average size:

- (20) Ik zoek een kamer.
 ‘I’m looking for a room.’
 Dat is geen kamertje, dat is een echte kamer.
 ‘That’s not a little room. It’s really a room.’

There are also privative oppositions and marked categories which involve adjectives, as in pairs like *oud* (‘old’) and *jong* (‘young’). At first glance, these two adjectives appear to stand in equal opposition to one another, *jong* meaning ‘having lived for only a short time’ and *oud* ‘having lived for a considerably long time’, but we soon come to realize that *oud* is also used as a neutral term, as the examples in (21) show:

- (21) Haar baby was pas drie dagen oud.
 ‘Her baby was only three days old.’
 Hoe oud ben je? Ik ben vier.
 ‘How old are you? I’m four.’
 Zij is één dag ouder dan hij.
 ‘She is a day older than he is.’

Jong is therefore the semantically marked member of the opposition *jong-oud*. The property that is present in *jong* but absent in *oud* is [not far advanced in life, youthful]:

- (22)
- | | | |
|--|------------|---|
| <i>oud</i> | <i>oud</i> | <i>jong</i> |
| ‘having lived a
considerably long time’ | | ‘having lived for
only a short time’ |
- ‘being of a certain age’

The question to be answered is the following: what exactly is the meaning of the unmarked members of a category?

According to Lyons (1977:308), *dog* can be a hyponym of itself, since it is sometimes in contrast with ‘bitch’ and sometimes superordinate to it (‘Is that dog a dog or a bitch?’). He continues: (this phenomenon) “is a direct consequence of semantic marking and should not be treated as an instance of polysemy”. Unfortunately, this is a mere statement, without argumentation; it is exactly the issue I would like to question. Do unmarked members of a category have one vague, indeterminate sense, or are they polysemous? The purely structural distinction equi-

pollent-privative opposition does not make reference to this issue, but I think it is actually crucial to understanding the meanings of unmarked categories.

In the case of *kamer* I don't think of polysemy at all – in spite of differences in size of the rooms in question, it has just one single sense –, but in the case of *oud* I do, based on sentences like the following:

(23) De oude en de jonge heer Jansen.

'The older and the younger Mister Jansen.'

Hun leraar was pas 23 jaar oud, maar de leerlingen vinden dat al oud.

'Their teacher was only 23 years old, but the pupils already think that's old.'

Hoe oud ben jij? Oh, 75, maar dat is tegenwoordig niet oud.

'How old are you? Oh, I'm 75, but nowadays that's not old.'

In my view, these different uses of *oud* do not correspond to a single sense, but to two different senses (see (22)).

Perhaps we can compare the meaning of *oud* with the meaning we find in words like those in (24), which show a broad sense in addition to their narrow sense:

(24)	narrow sense	broader sense
<i>dag</i> ('day')	time between sunrise and sunset	24 hours
<i>regen</i> ('rain')	condensed moisture	what comes down as rain: <i>een regen van kogels, confetti</i> ('a rain of bullets, confetti')
<i>zee</i> ('sea')	the salt water covering most of the earth's surface	an abundance, a great mass: <i>een zee van bloemen, mensen, vuur</i> ('a sea of flowers, people, fire')
French <i>fille</i>	daughter	girl

Regen, *zee* and *dag* have, in addition to a more specific meaning, a second, metaphoric or metonymic meaning which is broader. My hesi-

tations about the comparison with adjectives like *oud*, lies in the fact that this ‘broader’ meaning of, especially *regen* and *zee* is, unlike that of *oud*, at the same time the contextually restricted one, and so the ‘marked’ meaning. Accordingly, *rain* in a *rain of confetti* is not superordinate to *rain* ‘condensed moisture’.

A better example is the French noun *fille*, which can be used to refer to a daughter or a girl. As Goddard (2000:133) says, “any paraphrase which would be broad enough to include both kinds of use would have to say that *fille* meant something like ‘female human being’ and this would be too broad. Because we cannot find a single substitutable paraphrase, we have to posit two distinct meanings for *fille*”. This female personal noun example brings us to the meaning of non-feminine names.

4. The meaning of non-feminine names

4.1 Interpretation and meaning: vagueness and polysemy

How can we interpret Jakobson’s statement about the meaning of marked categories in general when it is applied to non-feminine names?

Whoever is not a woman must be a man, so the explicit presence of the sense-component [non-feminine] would imply the presence of the component [male]. But this is not the case. The unmarked category simply does not contain a gender-based component. When making use of this non-feminine meaning, we are usually dealing with a person who is non-feminine – that is, a man – though not exclusively. This meaning could also be used when one aims at a gender-neutral denomination.

We have seen examples in sentences such as (11) to (14). With *linguïst* in (11) we think of a woman, and in (14) we visualize a man, but this does not mean that the word *linguïst* contains the components [female] and [male], respectively, because we are guided by the use of the proper names *Lisa Cheng* and *Arie Verhagen* and the knowledge we have that is associated with these two names. We could easily account for these interpretations in terms of the non-gender-based meaning that is inherent in the unmarked word. In (12), *linguïst* is interpreted as ‘male linguist’ because of the context in which it stands in opposition to the feminine *linguïste*. So in this case we are indeed dealing with a specification of gender: it is not gender-neutral. The question to be asked is whether this interpretation can be accounted for in terms of the sense-component [non-feminine], the gender-neutral, ‘vague’ meaning, or

whether we should recognize that *linguïst* is polysemous, having two alternative meanings, namely ‘person’ and ‘male person’.

Let us compare this with the English noun *man*, which has two meanings: ‘male human being’ and ‘human being’; it is indeed polysemous, but what about the array of Dutch non-feminine names? Do *linguïst*, *dirigent* (‘conductor’), *loodgieter* (‘plumber’) and *hoogleraar* (‘professor’) have two meanings, namely ‘male person’ and ‘person’, or do they only contain the component ‘person’, a broad, somewhat vague meaning, which, when used in a particular context, is understood as ‘male person’?

Tuggy (1993) discusses various tests to be able to distinguish between vagueness and what he calls ambiguity. In (25) I have applied these tests to the nouns *kamer* and *regen* in order to differentiate between vagueness and polysemy:

- *The logical test: Can X and not-X be true?*

Dat is een regen van confetti maar geen regen

‘That is a rain of confetti but not rain.’

*Dat is een kamer met uitzicht, maar geen kamer.

*‘That is a room with a view, but not a room.’

- *The ‘linguistic’ test: grammatical constructions which are taken as requiring semantic identity: X does/did Z and so does/did Y. In case of ‘crossed’ readings without semantic oddness, the meaning of Z is vague, if zeugma results, Z is ambiguous.*

Jan heeft een zaal van een kamer, en Piet z’n hokje is ook een kamer.

‘Jan has got a room the size of a hall, and Piet’s little box is also a room.’

[Jan liep in de regen en Piet liep in een regen van confetti].

*Jan werd nat door de regen en Piet ook.

Jan was walking in the rain and Piet was walking in a rain of confetti.

* ‘Jan got wet with the rain, and Piet as well.’

Both tests indicate that *kamer* is vague while *regen* is polysemous.

What about, for instance, *linguïst*? The logical test indicates vagueness, because (26) is ill-formed:

- (26) *Zij is linguïst bij de opleiding Chinees, maar geen linguïst.
 ‘She is a linguist at the Chinese department, but not a linguist.’

The same is indicated by the linguistic test: there is nothing wrong with sentence (27) if we mean to say that the linguist already working there is a man, while leaving aside the gender of the newcomer:

- (27) De Opleiding Duits heeft één linguïst, maar ze krijgen een nieuwe.
 ‘The German department has one linguist, but they’re getting another one.’

My conclusion from these tests is that in accounting for the ways in which non-feminine names can be used, we can describe their meaning as being vague. The fact that we can use *linguïst* to refer to a man, a woman, and a person irrespective of his or her gender, does not imply that it is polysemous, i.e. that, in addition to the meaning ‘not-specifically feminine linguist’, it also means ‘male linguist’. There are, however, two factors which can cause us to interpret the word as containing the property [male]. First of all, there is the popular conception of lots of professionals, and a second factor is the actual way we make use of non-feminine names. It is these issues that we will address in sections 4.2 and 4.3.

4.2 Conception

If the unmarked names are gender-neutral, meaning ‘not-specifically female X’, where does the idea come from that they are not really neutral but, in fact, male? And why does it take us slightly longer to interpret sentence (7) as opposed to sentence (8)?

First of all, with many unmarked names we tend to think of men. Plumbers, boxers, captains, and so on, are men; that is to say, they are men in our conception or visualization of these words. They are associated with typically male professions, in which we find no or hardly any women. We could in this case speak of a *stereotype*:

- (28) “a socially determined collection of information associated with the extension of a word which a user must possess if he is assumed

to know the meaning of that word” (Geeraerts 1982:249, my translation).

Stereotype is a social concept, a kind of semantic norm. For many speakers the feature [male] is inherent in their knowledge of *plumber*. We could consider this to be knowledge of stereotypes, which forms part of the knowledge that the individual members of a speech community share, and as such it is social. Is this social knowledge contained in the meaning of the word?

Langacker (1987:154) defends this viewpoint in relation to the properties of (the word) *banana*, of which he mentions, among other things, the following aspects: a particular shape, color, taste, smell, and numerous other specifications like the knowledge that bananas are eaten, that they grow in bunches on trees, that they come from tropical areas and so on. He then proceeds by considering it “a pivotal problem of linguistic semantics: Which of these specifications belong to the meaning of the lexical item *banana* (...) Otherwise phrased, which of these specifications are linguistic (or semantic) in nature, and which are extralinguistic (pragmatic?) (...)”. My answer should hold no great suspense for the reader: All of these specifications are part of the meaning of *banana*. The distinction between semantics and pragmatics (or between linguistic and extralinguistic knowledge) is largely artificial, and the only viable conception of linguistic semantics is one that avoids such false dichotomies and is consequently *encyclopedic* in nature.”

According to Wierzbicka (1996), semantic knowledge *can* be separated from encyclopedic knowledge, but the difference with Langacker may largely be apparent. Wierzbicka, just like Langacker, considers various types of knowledge that structuralists would call encyclopedic to be part of the meaning of the word concerned. What is important is that she distinguishes between knowledge *of* and knowledge *about* a given concept, cultural knowledge deposited in language and other kinds of knowledge, whether scientific or non-scientific. In regard to the word *mouse*, there is, as she states, also knowledge “which is not part of the folk concept *reflected in language* – and a line can be drawn between that knowledge and the knowledge (and ideas) encapsulated *in* the word *mouse* itself” (Wierzbicka 1996:349).

A good Dutch example is the noun *beer* (‘bear’). Of course there is the ‘scientific’ knowledge that we have pertaining to bears – that they are dangerous, sometimes extremely aggressive animals – but the important

thing is that our knowledge is expressed through language, as is illustrated by the following Dutch expressions:

- (29) Zo sterk als een beer.
 ‘As strong as a bear.’
 Beresterk.
 ‘Bear-like strong’
 Een beer van een vent.
 ‘A man like a bear, i.e. a huge man.’

This knowledge, then, is indeed semantic knowledge. But how about the knowledge that bears also function as pets, and that not only children, but even grown-ups, have teddy bears? For many people, bears are not dangerous, but mollifying, and they don’t eat salmon, but honey. This social knowledge also belongs to our knowledge of the word *beer*, and given a word like *knuffelbeer* (‘cuddly bear’), following Wierzbicka, we could maintain that it is part of the meaning of *beer*. In fact, decisive for her distinction between semantic and encyclopedic knowledge, is the actual use of a word, a position similar to that of Langacker.

With some hesitation, therefore, I do think we may conclude that our cultural knowledge – that plumbers normally are men – is encapsulated in the word *plumber*, because this knowledge is reflected in language, the word *plumber* being almost always used in connection with men. This actual usage is the focus of the next section.

4.3 Usage

An important reason why we should consider the word *linguist* to be unmarked is that it can also be used to refer to a woman as well as to a person whose gender is irrelevant; but is that really what we do? Jakobson claims, as we have seen in the quotation in (18), that unmarked categories are used mainly to indicate the absence of the feature associated with the marked category. Is it indeed true that unmarked names are most often used in connection with men?

We can use unmarked names indicating a person’s position or occupation without having someone particular in mind, as in (30):

- (30) We zoeken een nieuwe directeur.
 ‘We are looking for a new director’.
 Een goede ambassadeur laat het landsbelang altijd voorgaan.
 ‘For a good ambassador the nation’s interest always comes first’.
 De manager van tegenwoordig kan niet meer zonder mobiele telefoon.
 ‘Today’s manager cannot do without a mobile phone’.
 Een goede advocaat is bijna niet te krijgen.
 ‘A good lawyer is hard to get’.

When used predicatively they are, true enough, gender-neutral, but their context of use may nevertheless assign gender, as in:

- (31) Mijn zwager is bedrijfsleider.
 ‘My brother-in-law is a manager’.
 Mevrouw Bakker, personeelschef bij de Leidse universiteit.
 ‘Mrs. Bakker, personnel manager at Leiden university’.

When we use them referentially, to indicate actual individuals, it becomes clear from the context whether we are dealing with a man or a woman:

- (32) Onze commandant zei dat *hij* er niet over piekerde ontslag te nemen.³
 ‘Our captain told us *he* would not even consider resigning’.
 Mijn therapeut kondigde aan dat *ze* drie weken op vakantie ging.
 ‘My therapist announced that *she* would be on holidays for three weeks’.

Sometimes gender is clear to the speaker but not to the listener:

³ This should not to be confused with the generic use of the personal pronoun *hij* and the possessive pronoun *zijn*, as in: ‘Van een pedagoog wordt verwacht dat *hij* tijdens *zijn* studie veertig uur per week werkt.’ (‘A pedagogue is supposed to work 40 hours a week during his study’.) In Dutch *hij/zijn* can be used as masculine and as gender-neutral.

- (33) Ik ken toevallig een specialist in plastische chirurgie
 ‘I happen to know a specialist in plastic surgery’.
 Onze fotograaf kwam een uur te laat
 ‘Our photographer was an hour late’.

This shows that although gender may not be inherent in the names itself, it is often assigned to it in actual language use. Unmarked names can be assigned male or female gender, depending on the context.

I have retrieved a number of Dutch names from a corpus, which comprises 27 million words taken from newspaper articles appearing in the *NRC-Handelsblad* in 1994 and 1995. My aim was to determine their frequency as well as their function. *Linguïst* is not a good example: it appears only twice, gender-neutrally; we more often come across *taalkundige*. Instead of *linguist* I will discuss *pedagoog* (‘pedagogue’), which is particularly interesting because, unlike *loodgieter* (‘plumber’), it does not necessarily make us think of a man: in Dutch society, it is a profession for both men and women, and as such not an isolated case. In fact, I believe the use of *pedagoog* to be illustrative of many names of professions.

Pedagoog appears 32 times, of which only once in a context denoting a woman, that is, if you are familiar with the referents of *Lea Dasberg* (a woman) and *De Levita* (a man):

- (34) Lea Dasberg heeft gelijk. Het meningsverschil tussen de psychiater en de *pedagoog* is een verschil in therapeutische opvatting. De Levita wil het zieke deel behandelen, zodat ...
 ‘Lea Dasberg is correct. The difference in opinion between the psychiatrist and the *pedagogue* is a difference in therapeutic view. De Levita wants to treat the diseased part in order to ...’

Reference to a man, on the other hand, is encountered 19 times, as in (35):

- (35) - De onderwijshervormer *Lighthart* leefde een eeuw geleden, een *pedagoog* in hart en nieren.
 ‘The educational reformer, *Lighthart*, lived a century ago, a *pedagogue* to the backbone.’
 - een ideale plek om *zijn* kwaliteiten als *acteur* en *pedagoog* in te zetten. Begin jaren tachtig hielp *hij* het Limburgse gezelschap ...
 ‘an ideal spot to make use of his qualities as an *actor* and *pedagogue*. In the early 80s *he* assisted the Limburgian company ...’
 - De *pedagoog*, zelf twee keer gescheiden, vier kinderen, vertelt over *zijn* eigen trauma.
 ‘The *pedagogue*, divorced twice, with four children, gives an account of *his* own trauma’.)
 - Veel later werd *hij* tenslotte *pedagoog* en psychotherapeut.
 ‘Much longer after that *he* eventually became a *pedagogue* and therapist’.

The gender-neutral usage was found only six times, as in:

- (36) Al krijgt hij wel een ochtend per week een *pedagoog* van de naburige school op bezoek.
 ‘Though he is visited one morning every week by a *pedagogue* connected to a neighboring school.’

In six cases usage was unclear. In (37) I summarize the actual use of the 32 instances of *pedagoog*:

(37)	<i>pedagoog</i>	32
	woman	1
	man	19
	gender neutral	6
	unclear	6

Pedagoog, then, is very frequently interpreted as referring to a man.

The female name *pedagoge*, which appeared 5 times, referred to a woman of course. The gender-neutral *opvoedkundige* appeared 3 times, once gender-neutrally and twice in relation to a woman. In (38) I give the answer to the onomasiological question which word do we use when we

want to speak of a woman, a man or a person who is a pedagogue by profession:

(38)	<i>pedagoog</i>	<i>pedagoge</i>	<i>opvoedkundige</i>
woman	1	5	2
man	19	-	
person	6	-	1

As we see, the word *pedagoog* is seldom used when we want to speak of a female pedagogue; in that case we prefer the word *pedagoge*. This strengthens the ‘male’ interpretation of *pedagoog*.

From (37) en (38) we can infer, that people have predominantly a male person in mind when the unmarked term *pedagoog* is used, even though it does not concern a typically male profession.

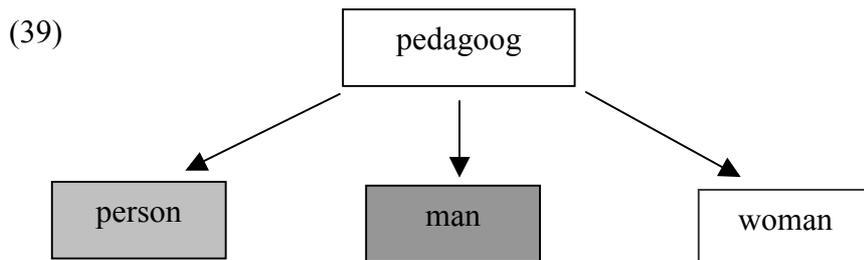
5. Conclusion

There are good reasons to assume that the meaning of non-feminine names can be described as ‘non-female person’. We have, however, seen that with many names we nevertheless picture a man and, furthermore, that other names are assigned male gender in actual language use. Jakobson’s claim holds in the case of *pedagogue*: usually it signifies a man, or at least the context makes us think of a man, but it is also used in a gender-neutral way, and occasionally, serves to denote a woman. Polysemy did not seem a satisfactory alternative to the meaning ‘non-feminine’, but how do we manage to account for both conception and usage?

I think that the best way in which this can be done is in terms of the so-called *Usage-based model*, which Langacker proposes in several publications. Central to this model are the actual use of the linguistic system and a speaker’s knowledge of this use. He regards the different meanings associated with a particular form as being a complex category of which the members “are analyzed as nodes in a network, linked to one another by various sorts of categorizing relationships” (1988:134). One such relationship is the “categorizing relationship that holds between a schema and a structure that ‘elaborates’ or ‘instantiates’ the schema” (134). We are dealing with a type of specialization: B is part of A though it is more precise, more detailed. As such, ‘rapid motion’ is the super-

schema of *to run*, of which ‘rapid 2-legged locomotion (person)’ – in addition to other meanings – is the prototypical meaning.

If we consider the names for professions, we could adopt the following schema: profession, followed by an X, allowing for the interpretations X is a man, X is a person, irrespective of gender, and X is a woman. These three options differ in salience and likelihood of activation as a result of usage and conception. I quote Langacker (2000: 36): “The nodes in such a network vary in their entrenchment and ease of activation.” The meaning of the non-feminine names for professions certainly does not exclude the interpretation according to which we are dealing with a woman, but this specification is considerably backgrounded, while [male] and [gender-neutral] are much more common. With many names for professions, the specification ‘man’ is much more prominent, or salient.⁴ I have tried to visualize this in (39):



As early as 1985, Bybee (1985:116) did not consider the question whether a particular form is or is not stored in the lexicon, to be a yes or no question, and proposed “to abandon this restricted, binary way of thinking about lexical storage, and treat the problem as the complex psychological problem that it is.” According to her, words differ from one another in lexical strength: “every time a speaker/hearer processes a word, it affects the lexicon by strengthening the representation of a lexical item”. It seems to me that this statement can be converted so as to be applicable to the status of the different possible uses or interpretations of lexical items, in this case names for professions: some are more entrenched than others.⁵

⁴ Cf. Cruse (2000:30) who refers to “richness and/or salience of conceptual content”.

⁵ In my opinion, this position comes close to that of Cruse (2000:50), whose

I can see two unquestionable advantages associated with an account of the meaning of non-female names that is, to a considerable extent, usage-based. To begin with, the question whether we are dealing with vagueness or polysemy is given no substantial importance. It is not a simple dichotomy. The various senses of a word do not have the same degree of salience; therefore, the question whether a certain meaning is conventional or not is likewise a matter of degree. The same holds for the question whether names for professions mean ‘man’ in addition to ‘person’.

The second advantage is that we can sufficiently account for potential changes in the use of words in general as well as in our names for professions. With *verpleger* we have seen that some non-feminine names are specifically male: we can only use *verpleegster* in relation to a woman. *Pedagoog*, on the other hand, *can* be used in reference to a woman, but as yet this is done sporadically in spite of the fact that quite a number of women are found in this profession. The reason for this is not that in that case we use *pedagoge* – although this might be a factor – but rather that at present we still tend to be talking about men in this profession rather than about women. There are hardly any female plumbers, which is why our conception of *plumber* is almost exclusively male. However, as conception and use of names of professions change, so does the extent to which the component [male] is salient.⁶

The meaning of non-feminine names of professions, like the meaning of all other words, is not fixed and invariable, but in permanent change. And it is not their meaning that dictates the possibilities to use them, but the other way around: actual use is decisive for the meaning of non-feminine names of professions.

References

- Adriaens, G. (1982). Vrouwelijke beroepsnamen in evolutie. *Forum der Letteren* 23, 1-17.
- Alphen, I. van (1983). Een vrouw een vrouw, een woord een woord. Over de gelijke behandeling van vrouwen en mannen en de consequenties daarvan voor beroepsbenamingen in het Nederlands. *Tijdschrift voor Vrouwenstudies* 14.4, 307-315.

view, in his own words, “imports a disturbing degree of fluidity into semantic structure”.

⁶ Cf. Adriaens (1982) and Gerritsen (2001, 2002) for changes in the use of Dutch professional terms.

- Bybee, J.L. (1985). *Morphology. A study of the relation between meaning and form*. John Benjamins, Amsterdam and Philadelphia.
- Caluwe, J. de & A. van Santen (2001). *Gezocht: functiebenamingen (m/v)*. Sdu Uitgevers, The Hague.
- Cruse, D.A. (2000). Aspects of the micro-structure of word meanings. In Y. Ravin & C. Leacock (eds.), *Polysemy. Theoretical and computational approaches*, 30-51. Oxford University Press, Oxford.
- Geeraerts, D. (1982). Prototypes en stereotypes. *Forum der Letteren* 23, 248-258.
- Gerritsen, M. (2001). Changes in professional terms in the Netherlands: Anglicisation and the neutralisation of gender. In T. van der Wouden & H. Broekhuis (eds.), *Linguistics in the Netherlands 2001*, 101-111. John Benjamins, Amsterdam and Philadelphia.
- Gerritsen, M. (2002) Language and gender in Netherlands Dutch: Towards a more gender-fair usage. In H. Bussmann & M. Hellinger (eds.), *Gender across Languages*, Volume II, 81-108. John Benjamins, Amsterdam and Philadelphia.
- Goddard, C. (2000). Polysemy: a problem of definition. In Y. Ravin & C. Leacock (eds.), *Polysemy. Theoretical and computational approaches*, 129-151. Oxford University Press, Oxford.
- Huisman, J. (1985). Sekse-neutrale persoonsnamen bestaan niet. *Onze Taal* 54, 70-71.
- Jakobson, R. (1957). *Shifters, Verbal Categories, and the Russian Verb*. Harvard University Press, Harvard.
- Langacker, R.W. (1987). *Foundations of Cognitive Grammar*. Volume I. Stanford University Press, Stanford.
- Langacker, R.W. (1988). A usage-based model. In B. Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, 127-161. John Benjamins, Amsterdam and Philadelphia.
- Langacker, R.W. (1999). A dynamic usage-based model. In M. Barlow & S. Kemmer (eds.), *Usage-based Models of Language*, 1-63. CSLI Publications, Stanford.
- Lyons, J. (1969). *Introduction to General Linguistics*. Cambridge University Press, Cambridge.
- Lyons, J. (1977). *Semantics*. Cambridge University Press, Cambridge.
- Tuggy, D. (1993). Ambiguity, polysemy, and vagueness. *Cognitive Linguistics* 4, 273-290.
- Wierzbicka, A. (1996). *Semantics. Primes and Universals*. Oxford University Press, Oxford.

The Dutch *way*

Arie Verhagen (Leiden University)

1. Introduction

When linguists describe the structure of a sentence, they typically use a rather restricted set of concepts, such as ‘accusative’, ‘subject’, ‘passive’, ‘verb phrase’, ‘goal’, and ‘adjunct’. These are quite abstract categories, that therefore have the advantage of being generally applicable. Similarly, teaching a system of linguistic analysis to students and teaching them how to apply it, is also typically restricted to abstract categories. The implicit assumption is that the structure of an object of linguistic analysis can be characterized both exhaustively and insightfully in terms of (combinations of) properties that this object as a whole as well as its parts share with *many* other elements in the language; notions with a limited range of applicability (e.g., going from still relatively abstract to quite concrete: ‘indirect object’, ‘benefactive dative’, ‘addressee (of a verb of communication)’, ‘promisee’) are considered less fundamental and derivative, so not really *required* for an adequate characterization of the grammatical structure of utterances in the language. This line of thinking follows the well-known – and in principle quite legitimate – idea that a scientific explanation of the properties of concrete, complex entities should be based on insights about the properties of the relatively simple (i.e. abstract) component parts of the complex entities, and the way they are put together.

A usage-based view of linguistic knowledge (Langacker 1988; Barlow & Kemmer 2000; Bybee & Hopper 2001) may be taken as challenging this principle of favoring abstract notions over more specific ones. The reason is the idea that the wide-ranging generalizations embodied in the most abstract notions will only emerge and get entrenched (to the extent that they can be used as productive rules) under very strong pressure of cumulative experience; in such a view it might therefore be expected that abstract notions will not often be sufficient for an exhaustive characterization of the grammatical structure of an utterance. This line of thinking follows more modern ideas about at least some sorts of complex systems, esp. living things, viz. that a scientific explanation

can never be complete without properties of (some) parts being seen as determined by the development of the system as a whole (hearts and livers, for example, cannot have evolved as independent ‘organs’ and they do not come into existence independently of an organism either, and organisms cannot ‘just’ be explained as assemblies of organs).

In this paper, I want to argue that in the case of language such a relatively radical interpretation of the usage-based view is precisely what is necessary in view of the facts. I will try to demonstrate this by looking at a set of phenomena in Dutch which formally involve the use of the word *weg* (‘way’) and semantically some notion of moving along a path (and spending energy in the process).¹ This is a rather limited class of phenomena, but a careful consideration of details of corpus data as well as linguistic (especially semantic) intuitions reveals that even at this level, at least three families of constructions have to be distinguished, on the basis of the fact that each of them exhibits some crucial properties that cannot be characterized in terms of general notions that apply to the set as a whole. The fact that this lack of generalizability is already observable at such a low level of grammatical organization, strongly suggests that (to put it paradoxically but succinctly) specific, small scale regularities are the rule rather than the exception. Such a view does raise the question, of course, how such a system of relatively independent constructions retains its coherence, and I will also make some suggestions in that respect.

2. How many ways?

Consider the following utterances.²

¹ The original reason for me to start looking into these phenomena was a remark by Goldberg (1996), based on personal communication from Annie Zaenen but certainly incorrect, that Dutch did not have an equivalent of the English *way* construction (see also Verhagen 2002).

² Unless indicated otherwise, the examples in this paper stem from the 1995 edition of the national newspaper *de Volkskrant* (available on CD-rom). For readability, I have sometimes changed subordinate clauses into main ones.

- (1) *In het gebouw kunnen bezoekers met computers hun weg zoeken.*
 In the building can visitors with computers their way seek
 ‘In the building, visitors can try to find their way with the help of computers.’
- (2) *Op de klanken van een tango zoeken eenlingen zich een weg door de nacht.*
 On the sounds of a tango seek loners REFL a way through the night
 ‘At the sounds of a tango, loners try to find their [themselves a] way through the night.’
- (3) *De priesters wurmen zich een weg door de gelovigen.*
 The priests squeeze REFL a way through the faithful
 ‘The priests squeeze their way through the faithful.’
- (4) *De weg naar de absolute top wordt voor vrouwen nog steeds geblokkeerd door een ‘glazen plafond’.*
 The way to the absolute top becomes for women yet always blocked by a glass-ADJ ceiling
 ‘For women, the way to the ultimate top is still blocked by a ‘ceiling of glass’.’

At first sight, all of these clauses have a noun phrase with the word *weg* as its lexical head functioning as a direct object. This noun phrase may be definite, marked with a possessive pronoun as in (1) or a definite article as in (4), or indefinite as in (2) and (3). The sentences apparently allow for a variety of verbs, so basically this also looks like a matter of free choice. Furthermore, there may be an indirect object (beneficiary), optional as usual, either marked with the preposition *voor* as in (4), or without a preposition (witness the ‘bare’ reflexive pronouns in (2) and (3)). Thus, there seem to be good reasons to consider these sentences simply different instantiations of the same underlying abstract pattern, essentially the basic pattern of transitive clauses, in each case just filled with different lexical material.

On a somewhat closer inspection, though, it soon becomes obvious that there are all kinds of ‘in-between’ regularities and constraints in the distribution of the features just mentioned, and that it is not at all a matter of ‘free choice’ of lexical material to be entered into an otherwise abstract formal syntactic pattern. For example, a possessive pronoun with

weg seems to be in complementary distribution with a reflexive beneficiary:

- (5)a. ?? *Bezoekers zoeken zich met computers hun weg door het gebouw.*
 Visitors seek REFL with computers their way through the building
- b. *Bezoekers zoeken zich met computers een weg door het gebouw.*
 Visitors seek REFL with computers a way through the building
 ‘Visitors try to find their way through the building with computers.’

And although the verb *zoeken* (‘to search, to seek’), witness (1) and (2), may be combined with either a reflexive or a possessive marking, the verb *wurmen* only occurs with reflexive marking:

- (3)’ ?? *De priesters wurmen hun weg door de gelovigen.*
 The priests squeeze their way through the faithful

In fact, there is a dependency between the type of marking of the beneficiary and (in)definiteness of the ‘direct object’: the presence of a reflexive beneficiary is incompatible with a definite article on *weg*.³

³ On the other hand, the presence of a prepositional beneficiary often seems to exclude an indefinite article:

- (i) *De Japanse bezetting maakte de/?? een weg vrij voor de communistische machtsovername.*
 The Japanese occupation made the/?? a way free for the communist take-over.
 ‘The occupation by Japan cleared the way for the communist take-over.’

But things may be somewhat less clear-cut here; see also section 3.2 and note 9.

- (2)' ?? *Op de klanken van een tango zoeken eenlingen zich*
 On the sounds of a tango seek loners REFL
de weg door de nacht.
 the way through the night

And while some verbs, such as *wurmen*, occur only with a reflexive beneficiary, the verb *vinden* ('to find') occurs only with a possessive marking of the direct object, although it is semantically closely related to *zoeken* ('to search, to seek'):

- (6)a. *Elke kunstenaar moet tussen deze twee polen zijn weg vinden.*
 Every artist must between these two poles his way find
 'Every artist will have to find his way between these two poles.'
 b. ?? *Elke kunstenaar moet zich tussen deze twee polen een weg vinden.*
 Every artist must REFL between these two poles a way find

Looking at actual usage, it is remarkable how dominant the weak form of the reflexive is. In the *Volkskrant* corpus, the full form *zichzelf* does not occur at all as a beneficiary in this type of sentences (i.e. in combinations with *weg*). A search in Dutch texts on the internet shows that although the strong form does show up sometimes, its frequency in this context is extremely low.⁴ In fact, speakers often have a problem accepting sentences of this type with *zichzelf*, such as (3)':

⁴ A search with Google (www.google.com) for the phrase 'baant zichzelf' resulted in 1 hit, as opposed to 795 hits for the phrase 'baant zich' (see below for the special role of the verb *banen*). A search for 'zichzelf een weg' produced 174 hits as opposed to 5800 hits for 'zich een weg', but 33 of these 174 were part of a prepositional phrase (while *zich* never occurred as such), and 52 were references to the same single line from a popular song (for the Dutch: *de eenzame fietser die kromgebogen over zijn stuur zichzelf een weg baant*), leaving 89 'true' cases of *zichzelf* in this context, which is a ratio of about 1:73 (*zichzelf* accounting for 1.3% of the total). This is very low compared to the

- (3)' ?? *De priesters wurmen zichzelf een weg door de gelovigen.*
 The priests squeeze themselves a way through the faithful

What this all suggests is that we do not have to do with different instantiations of the same underlying pattern, but actually with differences, both formal and semantic, in the underlying patterns, which therefore have to be relatively specific. I will now present a proposal as to the character of these different patterns, starting with the specific case characterized by the presence of a (weak) reflexive beneficiary.

3. The ways of Dutch

3.1. Making oneself a way

As the translations already indicate, the reflexive pattern exemplified by (2) and (3) constitutes the obvious Dutch translation equivalent of what is known as the *way* construction in English (Jackendoff 1990, Goldberg 1996). Examples of this construction are given in (7) and (8).

- (7) Pat pushed her way out of the room.
 (8) Volcanic material blasted its way to the surface.

The interesting thing about such sentences is that they share a number of systematic, correlated properties in both form and interpretation which cannot be explained on the basis of the formal and semantic features of the words and the general grammatical structure of the sentences. Specifically, the subject referent creates a (possibly metaphorical) path and/or removes obstacles on it, and travels it, while a verb like *push* normally neither indicates movement of the subject nor the creation of something. Moreover, the presence of the noun *way*, marked with a possessive pronoun, is a necessary condition for this interpretation (cf. Jackendoff 1990 and Goldberg 1996, and the references cited there, for more details). Thus, one has to conclude that a syntactic pattern that may

average ratio (in web-pages containing the word *willekeurig* found by Google) of about 1:3 (*zichzelf* accounting for 24% of the total).

roughly be indicated as ‘to *verb* one’s way +locational adjunct’ is itself conventionally associated with a specific meaning, and stored in long-term memory of language users; a form-meaning pairing of this type is called a construction. Using a notation taken from Goldberg, the *way* construction may be represented as follows:

- (9)
$$\left[\begin{array}{cccc} \text{Sem:} & \text{creator,} & \text{create-move,} & \text{created-way,} & \text{path} \\ & | & | \text{means} & | & | \\ \text{Syn:} & [\text{SUBJ}_i & [\text{V} & [\text{POSS}_i \textit{way}] & \text{OBL}]] \end{array} \right]$$

The bottom line (‘Syn’) mentions the obligatory elements of the syntactic pattern that characterizes the construction, and the top line (‘Sem’) mentions the components of its meaning. The connecting vertical lines indicate which components are associated with which elements, and what (if any) features of interpretation are imposed upon certain elements, in this case: the fact that the process mentioned by the verb (such as the pushing in (7)) is interpreted as the means by which the path is created and/or traveled.

Consider now the Dutch sentences (2) and (3), and some other similar instances of the same pattern in (10) and (11).

- (10) *Zo bluffte zij zich een weg uit Auschwitz.*
 Thus bluffed she REFL a way out-of Auschwitz
 ‘That was the way she bluffed her way out of Auschwitz.’
- (11) *Twee bussen boren zich een weg naar het hart van Istanbul.*
 Two buses drill REFL a way to the heart of Istanbul
 ‘Two buses are drilling their way to the heart of Istanbul.’

The similarities with the English *way* construction are obvious. There is a constant lexical element *weg*, a variety of verbs indicating the means by which a path is created, and prepositional phrases specifying the path being traveled. The lexical meanings of the verbs in the construction do not have to contain a component of movement (cf. *bluffen* and *boren* in (10) and (11), respectively), but the referents of their subjects all move, clearly because of the meaning of the construction itself; in fact, the transitivity of (10) can clearly only be attributed to the construction (*bluffen* is not itself a transitive verb). This all provides very good

reasons to consider this pattern the Dutch analog of the *way* construction. However, there are also differences. I will first simply describe the differences; possible theoretical consequences will be discussed in the final section, after we have analyzed more members of the same family.

To begin with, there is a noticeable difference in the syntax. Whereas the relationship between the subject and the created way is marked by a possessive determiner in English, it is marked with a weak reflexive (*zich*) in indirect object position in Dutch. The representation of the Dutch construction may thus be given as in (12):

$$(12) \left[\begin{array}{ccccccc} \text{Sem:} & \text{creator,} & \text{create-move,} & \text{for-self,} & \text{created-way,} & \text{path} & \\ & | & | \text{means} & | & | & | & \\ \text{Syn:} & [\text{SUBJ}_i & [\text{V} & [\text{REFL}_i & [\text{een weg}] & \text{OBL}]] & \end{array} \right]$$

In itself, it may not be immediately clear whether this difference is theoretically significant or not; I will get back to this question in the final section. But in any case, the fact that the ‘choice’ of a reflexive marking does not seem derivable from more general properties of Dutch (since a possessive marking would have been perfectly possible), is yet another argument for the hypothesis that the construction is itself stored in long term memory, as a conventional unit.

Another difference concerns the verbs used. Table 1 lists the verbs occurring in the Dutch construction.⁵

⁵ The absolute numbers are different from those in Table 1 in Verhagen (2002: 412). This is due to the possibility of using a more advanced search algorithm for this study, allowing the inclusion of variables between the elements ‘*zich*’ and ‘*een weg*’ in the search expression. On the whole, however, the proportions in the frequencies are the same as in the previous study, so the conclusions on this point are not changed.

Tokens/verb	Verbs	Total number
1	<i>beitelen</i> ('chisel'), <i>boren</i> ('drill'), <i>graven</i> ('dig'), <i>knagen</i> ('gnaw'), <i>knippen</i> ('cut out'), <i>ploegen</i> ('plough'), <i>schermen</i> ('fence'), <i>schieten</i> ('shoot'), <i>verschaffen</i> ('provide'), <i>wurmen</i> ('wiggle')	10
2	<i>bluffen</i> ('bluff'), <i>kronkelen</i> ('twist'), <i>vreten</i> ('eat, gnaw')	6
3	<i>slaan</i> ('hit'), <i>snijden</i> ('cut')	6
4	<i>zoeken</i> ('search, seek')	4
7	<i>vechten</i> ('fight')	7
59	<i>banen</i> (?)	59
	18	92

Table 1. *Verbs used in the Dutch (reflexive) 'way' construction (Volkskrant 1995)*

The pattern is obvious. In a considerable number of instances, a verb is used that contributes its lexical meaning to the interpretation of the sentence, as the means by which the path is created/traveled. But in the majority of cases just a single verb occurs, namely *banen*. Some examples are:

- (13) *De spermasliertjes trachten zich een weg naar het eitje te banen.*
 The sperm-strings-DIM try REFL a way to the egg-DIM to 'banen'
 'The strings of sperm try to make their way to the egg.'
- (14) *Twee figuren in zwart pak banen zich met grote zwemvliezen aan hun voeten een weg door de menigte.*
 Two figures in black suit 'banen REFL with big flip-pers on their feet a way through the crowd
 'Two individuals in black suits make their way through the crowd with large flippers on their feet.'

It is immediately clear that this is the default verb for the construction, but it is less obvious what its role in the language is, i.e. what it means. When asked, speakers of Dutch tend to answer: 'to make, namely a way'.

The real point is that it only occurs in this kind of construction;⁶ it actually does not have a meaning independently of a combination with *weg*, so if one tries to describe its meaning, then one essentially ends up with something very similar to the meaning of the entire construction. The role of this verb is to make it possible to actualize the meaning of the construction *without* adding information about the specific manner in which the path is created or traveled, as is the case when another verb than *banen* is used. English also has a verb with this function in the *way* construction, viz. *to make*. So whereas English has, so to speak, opted for the strategy of using a verb with such a general meaning that it exactly fits the role of the verb slot in the construction (cf. (9)), Dutch employs a verb that is highly specific for the construction for the same purpose. This observation allows us to specify rather precisely to what extent the two languages differ at this point: they share the feature that the constructions have a prototype, the use of which simply realizes the construction without adding anything to its meaning, but they differ in the choice of a general vs. a specific verb in this prototype.

An aspect in which English and Dutch appear not to differ at all is the character of the oblique path phrases, specified in the representations (9) and (12). Table 2 contains the path-markers, i.e. prepositions and adverbs, found in the corpus.⁷

⁶ It has not always been like this. See Kramer (2002) and Verhagen (2002) for an overview of the origin and the development of the Dutch *way* construction.

⁷ The remark in note 5 also applies to Table 2. There are a few sentences with two (partial) path-phrases (e.g. ‘over and preferably through X’), whence the total number of cases in Table 2 is higher than in Table 1. An example is:

- (i) *Het hete gas had zich een weg weten te banen door*
 The hot gas had REFL a way know to ‘banen’ through
de laag kit naar de binnenste van de twee O-ringen.
 the layer cement to the innermost of the two O-rings
 ‘The hot gas had managed to make its way through the cement layer
 to the innermost one of the two O-rings.’

Tokens/marker	Preposition or adverb	Total number
	Absent	4
1	<i>in</i> ('in(to)'), <i>ondergronds</i> , ('underground'), <i>op</i> ('on'), <i>tot</i> ('till'), <i>uit</i> ('out of'), <i>via</i>	6
3	<i>langs</i> (2 'past'; 1 'along'), <i>over</i> ('over'), <i>terug</i> ('back')*	9
5	<i>tussen</i> ('(in) between')	5
26	<i>naar</i> ('to')	26
55	<i>door</i> ('through')	55
	12	105

*all three cooccurring with the verb *vechten* ('to fight')

Table 2. OBL-markers used in the Dutch 'way' construction
(*Volkskrant* 1995)

The preposition *door* ('through'), occurs in about 50% of the cases. While adjunct phrases are in general considered optional, oblique phrases are very frequent here, so they appear to be an integral part of the conventional schema. In fact, a close look at the four instances where a path phrase is missing confirms this; two cases contain phrases marked with *in*, which, though they do not themselves denote a path, refer to an area containing obstacles to be avoided or removed. In the remaining two, the nature of the path can straightforwardly be inferred from the immediate context; consider (15), for example.

- (15) *Ik zet een hek om mijn veldje. Maar dat konijn*
I put a fence around my field-DIM. But that rabbit
graaft zich een weg en eet door.
digs REFL a way and eats through
'I put a fence around my little field. But this rabbit digs its way
and continues to eat.'

The first sentence makes it absolutely clear what the path is that the rabbit creates and travels: from the outside into the little field, the crops of which the speaker is desperately trying to protect.

All in all, in view of the use of verbs as well as of path markers in the corpus, there is ample reason to assume that speakers of Dutch have

stored in their memory both the highly specific schema *zich een weg banen door X* (with both the verb and the path-marker lexically specified), as well as the more general, superordinate schema (*zich een weg +V + OBL*) that we have been considering so far (cf. (12)); the specific schema is, as already suggested above, the prototype of the general one. Other specific schema's might also be stored independently, albeit with less degree of entrenchment as the prototype; in view of the correlation between the adverb *terug* ('back') and the verb *vechten* ('to fight'), the combination *zich een weg terug vechten* seems a possible candidate. Assuming this to be the case, the relations between the patterns can be represented as in the partial taxonomic network below.

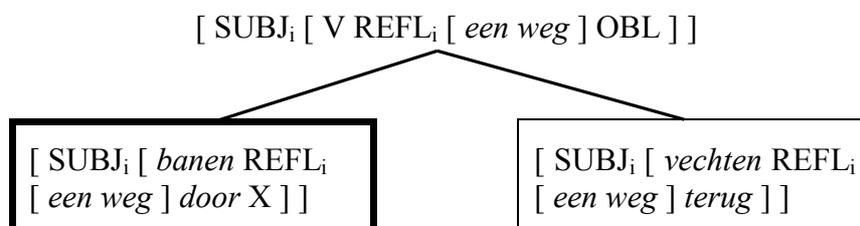


Figure 1. *Dutch way construction network (partial)*

It is an important and intriguing question how this partial network is integrated into the network of constructions in Dutch in general, and how this compares to the situation in English; I will get back to this in the final section.

A final point relating to syntax is the following. According to Jackendoff (1990) and Goldberg (1996), as well as others, the verb in the *way* construction does not always have to be interpreted as indicating the means by which a path was created; instead, it may simply describe some activity *accompanying* the movement along a path. This usage does not exist at all in Dutch. Thus (16) is OK for (at least some) speakers of English, with no need to impose the reading that the whistling was the instrument for removing obstacles, but (17) in Dutch can only mean that he created a way to the front door by whistling, and hence it is very strange.

(16) He whistled his way to the front door.

- (17) ?? *Hij floot zich een weg naar de voordeur.*
 He whistled REFL a way to the front-door

This semantic difference seems to be connected rather directly to the difference in syntax. Israel (1996) shows that the modern English construction is a case of a diachronic ‘blending’ of two constructions. The original situation was that on the one hand, there was a construction in which transitive verbs (of creation) occurred with the NP *one’s way* as direct object (of the type *He made/paved his way*), and on the other hand a construction in which verbs of movement occurred with the NP *one’s way* in an adverbial role (of the type *He went his way*). In both patterns, the sets of verbs that could be used were extended through analogy, until such a degree of overlap between the two classes emerged that many language users interpreted specific cases as instantiations of the same pattern, which allowed for more than one meaning.

From the point of view of Dutch, it seems clear that if the parallel linear structure of both patterns (*to make one’s way – to go one’s way*) has not actually promoted this development, it has in any case not prevented it. There are expressions with *weg* in Dutch that indicate ‘movement along a path’, but they are structurally more different from the *way* construction than in English, if only because there is no (pronominal) indirect object:

- (18) *Hij ging zijns weegs.*
 He went his-GEN way-GEN
 ‘He went his way.’
- (19) *Zij vervolgde haar weg.*
 She continued her way
 ‘She continued on her way.’

Thus, the fact that in Dutch the semantic difference is correlated with a clear syntactic difference contributes to understanding why the modern Dutch *way* construction is not polysemous in the same way as its English counterpart, and lacks the possibility of an ‘accompanying activity’ reading. Some other, arguably more important, instances of possessive-marked expressions with *weg* will be discussed in section 3.3.

3.2. Easing and blocking the way

In the preceding section, it was established that the Dutch pattern *zich een weg +V +OBL*, is associated with a specific meaning, which is highly, though not totally, comparable to the meaning of the English *way* construction, as was in fact suggested by the similarity between the representations (12) and (9). The fact that a detailed analysis shows that the similarity is neither formally nor conceptually perfect, is yet another indication of the conventional, symbolic nature of constructions. Having established that the specific pattern with a reflexive ‘beneficiary’ (*zich*) and indefinite ‘direct object’ (*een weg*) has this particular meaning, we can now proceed to ask to what extent the more or less different patterns mentioned in section 2 also have this meaning. The first candidate that should be considered is the one in which the ‘beneficiary’ is not a bare reflexive pronoun, but a nominal marked with the preposition *voor*, i.e. sentences of the following type:⁸

- (20) *Sex baande voor hem ook de weg naar de roem.*
 Sex ‘banen’-PAST-SG for him also the way to the fame
 ‘Sex also paved the way to fame for him.’
- (21) *Dit koor baande de weg voor kleinere ensembles.*
 This choir ‘banen’-PAST-SG the way for smaller ensembles
 ‘This choir paved the way for smaller ensembles.’

As the examples show, the verb *banen* also occurs in this pattern, which may reinforce the idea that it would be entirely possible that the only difference between such cases and the ones discussed in section 3.1 is the

⁸ The material discussed in this section was collected by means of putting together the results from a number of specific searches through the corpus. While the presence of a reflexive pronoun allowed a relatively specific search for examples of the reflexive pattern, a general search pattern for the non-reflexive pattern could only contain the phrases ‘de weg’ and ‘een weg’, which resulted in too many irrelevant hits. Instead, a number of specific searches were performed, combining these phrases with (instances of) *banen*, and (in view of the results from the first search) also with one of the prepositions *voor* and *naar*. Although this provides a sufficient basis for the claims to be made in the remainder of this paper, this is the reason that I cannot give tables of all verbs and all oblique markings used with the non-reflexive pattern.

non-reflexivity of the relation between agent and ‘beneficiary’; in that case, they should actually both be assigned to the same slightly more abstract pattern: +NP [*DET weg*] +V, hypothetically meaning: ‘to create a/the path for NP and have NP travel it’. However, a first indication that such a reduction to one abstract pattern is unwarranted, is the fact that the ‘direct object’ in the cases of section 3.1 is always indefinite, while it is definite in (typical) examples such as (20) and (21). In fact, no combinations of *zich* and definite *de weg* were found in the present corpus, whereas in nearly all non-reflexive clauses *weg* is marked as indefinite,⁹ and many speakers actually find indefiniteness obligatory in such cases; this is a problem for the idea of an abstract pattern, as this would in principle predict free variation in this area.¹⁰ But even more telling evidence is provided by semantic considerations. These may be demonstrated on the basis of the examples given above and the following ones:

- (22) *Daarmee opent hij de weg naar machtsmisbruik.*
 Therewith opens he the way to power-abuse
 ‘With that he is opening the way to abuse of power.’
- (23) *Zijn concessie maakte de weg vrij voor ondertekening van het akkoord.*
 His concession made the way free for signing of
 the agreement
 ‘His concession cleared the way for the agreement to be signed.’

⁹ Despite the fact that these do occur occasionally: Kramer (2002) found a few examples in her 20th century material. Further research will have to show whether the dependency between definiteness of *weg* and non-reflexive beneficiaries is really weaker than that between indefiniteness and reflexive ones.

¹⁰ This is actually another demonstration of the insight that ‘alternations’ suggested by paraphrase relations (in this case: that *zich een weg banen*, ‘to make one’s way’, can be paraphrased as *een weg voor zichzelf banen*, ‘to make a way for oneself’) are in fact very weak generalizations (cf. Goldberg 2002). See also p. 54 in section 4 below.

- (24) *Deze uitspraak effent de weg voor de scheiding van de carrières van rechters en aanklagers.*
 This decision levels the way for the separation of the careers of judges and prosecutors
 ‘This decision paves the way for the separation of the careers of judges and prosecutors.’
- (25) *Hij liet de weg voor onderhandelingen open.*
 He let the way for negotiations open
 ‘He left the way for negotiations open.’
- (26) *Ze blokkeerden de weg tot de kassa's voor de rest van de menigte.*
 They blocked the way to the cash-registers for the rest of the crowd
 ‘They blocked the way to the cash-registers for the rest of the crowd.’

For one thing, it seems doubtful whether the role of ‘beneficiary’ is at all obligatory in this pattern. Recall that I argued above, in connection with examples such as (15), that in reflexive cases, this role was an obligatory part of the interpretation even when it was not expressed. In the non-reflexive pattern, there is also always an oblique phrase, but it is not necessarily marked with the preposition *voor* (as would have to be the case if the required role was that of beneficiary): there are quite a number of examples in which the only prepositional phrase present is marked with *naar* (‘to’, ‘towards’) as in (22), i.e. a directional one indicating a part of the relevant path. The ‘beneficiary’ phrases ‘for him’ and ‘for the rest of the crowd’ may also be left out from (20) and (26), respectively, without the sentences becoming unacceptable or changing their general meaning. Moreover, in several of the cases in which the preposition *voor* is used, the role of the participant involved is arguably not at all that of a beneficiary, but a goal or endpoint, e.g. the signing of the agreement in (23) and the separation of careers in (24); in fact, one might argue that even the roles of the smaller ensembles in (21) and the negotiations in (25) are better characterized as goal than as beneficiary. Thus we have to conclude that what is required in these cases is the presence of some element with a goal-like role, but it does not have to be specifically a beneficiary (notice that the concept ‘goal’ seems to be part of that of beneficiary). Language users may occasionally still want to mark both a specific beneficiary and a specific goal (as in (20) and (26)), but the point

is that the role of beneficiary is not *obligatory* in the non-reflexive pattern, while it is in the reflexive one.

The second, and most revealing semantic observation concerns the verbs and their semantic relationship with the construction as a whole. The verbs occurring in the reflexive pattern may be both transitive and intransitive, and when used in this context, they indicate activities that are instrumental in creating and/or travelling the path involved. But in the present cases, the set of verbs is much more restricted, with a much more specific semantic profile. They are transitive verbs (or verbal compounds) meaning ‘to make open/free’, ‘to leave open/free’ or ‘to block’, notions connected directly to the concept of a barrier (*viz.* (not) creating or (not) removing it). What we have here is exhaustively characterized as the dimensions of the conceptual domain of Force Dynamics (Talmy 1988), which also plays a crucial role in the semantics of causative constructions (Verhagen & Kemmer 1997; Loewenthal, this volume). The different paradigms of verbs used in the two patterns suggest an important semantic difference. The reflexive one implies that a barrier is removed and the relevant path is actually traveled, but the present one only ‘raises the issue’ of a barrier; it may simply express that travelling the path is permitted (because no barrier is created, as in (25)), without the implication that the endpoint is actually reached.¹¹ It may even express that travelling the path is *prevented* (a barrier is created, as in examples (4) and (26)). Notice that the verbs *vrijlaten* (‘to leave free’) and *blokkeren* (‘to block’) cannot be used in the reflexive construction:

¹¹ Notice that something similar holds for the English phrase *to pave the way* (contrary to *to make one’s way*, with possessively marked *way*). The following example, from the table of contents of *Scientific American* vol. 285, number 1 (July 2001), provides a nice illustration:

(i) **Frozen light**

Halting photons paves the way for quantum computing and tabletop black holes.

The goals of quantum-computing en tabletop black holes are, of course, not yet realized by the halting of photons.

- (27) ?? *Hij liet zich een weg vrij naar een andere baan.*
 He let REFL a way free to an other job
 [Supposed reading:
 ‘He left the way to another job free for himself.’]
- (28) ?? *Zo blokkeer je je een weg naar de top.*
 Thus block you you a way to the top
 [Supposed reading:
 ‘In that way, you will block a way to the top for yourself.’]

On the other hand, *intransitive* verbs that indicate activities instrumental in creating and travelling a path – which can characteristically be used in the reflexive construction – precisely cannot be used with the non-reflexive pattern:

- (29) ?? *Hij vocht/blufte (voor haar) de weg naar de troon.*
 He fought/bluffed (for her) the way to the throne.
 [Supposed reading: ‘He cleared the way to the throne (for her) by fighting/bluffing.’]

The difference between the roles of the verbs in the two patterns also implies a difference in the roles of the subjects: while this is a volitional agent (who has a goal to reach) in the reflexive construction, it is more of a Source-of-Energy in the cases considered here.

In fact, the conventional association of the reflexive pattern with the accomplishment of a goal holds not so much at the level of propositional content, but at the speech act level; the idea of reaching the endpoint of a path is not just evoked, but actually *asserted*. Thus, instances of the Dutch *way* construction cannot be negated in a straightforward manner:

- (30) ?? *Zij baanden zich geen weg door de menigte.*
 They ‘banen’-PAST-PLUR REFL no way through the crowd
 [Supposed reading:
 ‘They did not make their way through the crowd.’]

- (31) ?? *De bussen boorden zich niet een weg naar het hart van de stad.*
 The buses drilled REFL not a way to the heart of the town
 [Supposed reading:
 ‘The buses did not drill their way to the heart of town.’]

The only combinations of this construction with negation actually found involve some notion of modality, specifically issues of ability or desire:

- (32) *Je kunt je geen weg schieten naar het hart van een volk.*
 You can you no way shoot to the heart of a people
 ‘You can’t shoot your way to the heart of a nation.’

But negation is not odd with unmodified instances of the non-reflexive pattern:

- (25)’ *Hij liet de weg voor onderhandelingen niet open.*
 He let the way for negotiations not open
 ‘He did not leave the way for negotiations open.’

Although they are not very frequent, the following corpus examples show that this possibility is realized for the non-reflexive pattern (unlike negation of the reflexive construction), both for the ‘clearing’ and the ‘blocking’ type of instances:

- (33) *Een promotie verlicht duidelijk niet de weg naar het grote geld.*
 A getting-Ph.D. eases clearly not the way to the big money
 ‘Having a doctoral degree clearly does not ease the way to big money.’
- (34) *Hopelijk verspert dit niet de weg van betere Franse films tot de Nederlandse bioscopen.*
 Hopefully bars this not the way of better French movies to the Dutch cinemas
 ‘Hopefully, this does not bar the way of better French movies to Dutch cinemas.’

All in all, we have now established both differences of form as well as of conceptual content between the reflexive and non-reflexive patterns, so that the conclusion should clearly be that each of these constitutes an independent symbolic unit, a conventional link of form and function with internal structure, i.e. a construction. We may represent the second construction as in (35):

$$(35) \left[\begin{array}{cccc} \text{Sem:} & \text{source, force dynamics,} & \text{way,} & \text{goal} \\ & | & | & | \\ \text{Syn:} & [\text{SUBJ}] & [\text{V}] & [\text{de weg}] \text{ OBL } \end{array} \right]$$

It is obviously related, both formally and conceptually, to the reflexive Dutch *way* construction discussed in 3.1; (12) is repeated here for convenience:

$$(12) \left[\begin{array}{cccc} \text{Sem:} & \text{creator, create-move, for-self, created-way, path} \\ & | & | \text{ means} & | \\ \text{Syn:} & [\text{SUBJ}_i] & [\text{V}] & [\text{REFL}_i] \text{ [een weg] OBL } \end{array} \right]$$

At the same time, it is clear that neither one is an instance of the other, and that they also cannot be reduced to a single common pattern, as they participate in partly distinct relations of (dis)similarity. In particular, the sets of verbs allowed in the V-slots of each pattern are different, except for the special verb *banen* which is idiosyncratic in being the default verb for precisely these two constructions. I will discuss the theoretical consequences of this situation in the final section.

3.3 Finding one's way

The independence of the reflexive and non-reflexive *way* constructions, justified on empirical grounds in the previous sections, explains several of the restrictions and dependencies noted in section 2. One major issue remains, viz. the occurrence of 'English-like' possessively marked phrases with *weg*, as in (1):

- (1) *In het gebouw kunnen bezoekers met computers hun weg zoeken.*
 In the building can visitors with computers their way seek
 ‘In the building, visitors can try to find their way with the help of computers.’

I will only briefly indicate the main reasons why this is an instance of yet another independent specific, albeit related, pattern. In view of the preceding discussion, it will come as no surprise that the verbs occurring in it and their semantic relation to the pattern as a whole, again constitute the essential clue. Table 3 provides an overview of these verbs.

Tokens/verb	Verbs	Total number
1	<i>aanvangen</i> (‘start’), <i>forceren</i> (‘force (open)’), <i>kennen</i> (‘know’), <i>struikelen</i> (‘stumble’), <i>vechten</i> (‘fight’), <i>voortzetten</i> (‘pursue’), <i>weten</i> (‘know’)	7
2	<i>kiezen</i> (‘choose’), <i>uitstippelen</i> (‘map out’)	4
12	<i>vervolgen</i> (‘continue on’)	12
24	<i>zoeken</i> (‘search, seek’)	24
107	<i>vinden</i> (‘find’)	107
	12	154

Table 3. *Verbs used with ‘POSS-way’ (Volkskrant 1995)*

The first thing to be noticed is the absence of *banen* from this table;¹² the prototypical verb here is *vinden* (‘to find’); its relative frequency in this

¹² It cannot be entirely excluded that this is partly due to the corpus used for this study (*de Volkskrant*). The verb *banen* did occur rather regularly with possessive-marked *weg* in older varieties of Dutch, and it still occurs occasionally, as a search (with Google) on the internet for the patterns *zijn weg baant/baant zijn weg* showed. It is clear, however, that this is at best a small minority pattern: the number of hits for *baant zijn weg* was about 20, while for *vindt zijn weg*, *zoekt zijn weg* and *vervolgt zijn weg*, it was about 570, 200, and 300, respectively. For comparison: the number of hits for the reflexive *baant zich een weg* was about 470.

corpus (over 69%) is even higher than that of *banen* in the reflexive *way* construction (64%). Two typical examples are (36) and (37).

- (36) *Veel kunst vindt via vlooienmarkten zijn weg naar de kopers.*
 Much art finds via flea-markets his way to the buyers
 ‘A lot of art finds its way to the buyers via flea markets.’
- (37) *In dit spanningsveld heb ik getracht mijn weg te vinden.*
 In this field-of-tension have I attempted my way to find
 ‘In this field of tension, I have tried to find my way.’

In view of the very close semantic relationship between *vinden* (‘to find’) and *zoeken* (‘to search, to seek’), one might actually want to include the latter in the prototype. It is worth noting that the ratio of *vechten* (‘to fight’, the next most frequent verb in the reflexive *way* construction) to *banen* is 1 : 8.5, whereas the ratio of *zoeken* to *vinden* in the possessive pattern is 1 : 4.5. The two most frequent (semantically related) verbs in the possessive construction account for 85% of all the instances, while the two most frequent (semantically unrelated) verbs in the reflexive construction account for 71%. Clearly, the semantic variation in the verb slots in the possessive pattern is much more restricted than in the reflexive pattern. What appears to be common to the instantiations of this pattern is that there is not really a force creating a path, but some form of motion, ranging from actual movement along the path (in the most typical cases) to the purely mental movement, i.e. scanning, involved in ‘knowing’ and ‘mapping out’ one’s way; in some cases, especially with the verb *zoeken*, they may be involved simultaneously, as in (1) and in the following example:

- (38) *Nina en Vladimir zoeken hun weg tussen de ruïnes van*
 Nina and Vladimir seek their way among the ruins of
hun stad.
 their city
 ‘Nina and Vladimir try to find their way among the ruins of
 their city.’

In view of the fact that the path is not conceptualized as being created, the use of the possessive marking, and hence definiteness of the *weg*-phrase, can be seen as motivated: the path exists independently of the present event. In fact, in many cases the path referred to may be inter-

preted as in some sense ‘inherent’ to the mover, as specifying a ‘teleological’ quality of the mover (the notion of consumer is part of the knowledge of the purpose of goods, etc.).

In two cases in this material, a verb is used whose lexical meaning imposes the reading of an activity of attempting to reach a goal, suggesting the creation of a path (*forceren*, ‘to force (open)’, *vechten*, ‘to fight’). This number is so low that they could be taken as errors but also as analogical extensions (removing obstacles not being incompatible with movement); in any case, this is clearly not a well-entrenched sub-pattern.

The possessive pattern also requires an oblique phrase, but it does not have to specify (a part of) the path traveled or scanned. Sometimes it does (e.g. (36)), but it may also indicate the region within the boundaries of which the path is located. All in all, I propose the following representation for this third type of *way* construction in Dutch:

(39)	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding-right: 1em;">Sem:</td> <td style="padding-right: 1em;">mover/ scanner,</td> <td style="padding-right: 1em;">(mental) motion,</td> <td style="padding-right: 1em;">way,</td> <td style="padding-right: 1em;">location</td> </tr> <tr> <td style="padding-right: 1em;">Syn:</td> <td style="padding-right: 1em;">[SUBJ_i]</td> <td style="padding-right: 1em;">[V]</td> <td style="padding-right: 1em;">[POSS_i <i>weg</i>]</td> <td style="padding-right: 1em;">[OBL]]</td> </tr> </table>	Sem:	mover/ scanner,	(mental) motion,	way,	location	Syn:	[SUBJ _i]	[V]	[POSS _i <i>weg</i>]	[OBL]]
Sem:	mover/ scanner,	(mental) motion,	way,	location							
Syn:	[SUBJ _i]	[V]	[POSS _i <i>weg</i>]	[OBL]]							

The only verb from Table 3 (besides the two single cases of *forceren* and *vechten* mentioned above) that seems fully compatible with both this construction and the reflexive *way* construction (12), is *zoeken*: Table 1 shows that it occurs 4 times in that pattern; an example is (2), repeated below.

- (2) *Op de klanken van een tango zoeken eenlingen zich een
 On the sounds of a tango seek loners REFL a
 weg door de nacht.
 way through the night
 ‘At the sounds of a tango, loners try to find their [themselves a]
 way through the night.’*

The fact that it is this verb and not the prototype *vinden* which easily occurs in both patterns, should not come as a surprise: unlike *vinden*, the lexical meaning of *zoeken* implies effort on the part of the subject referent. It thus fits the semantics of the reflexive *way* construction rather

well, and the creation-reading can be imposed on the verb by the construction in many other cases, too. So we have now provided an analysis of the last of the remaining problems from section 2.

4. Conclusion: the organization of a grammar

At a number of points in the preceding sections, I explicitly emphasized that the three constructions discussed are truly independent units that cannot be reduced to each other or to completely general rules of grammar, as they each exhibit specific, irreducible properties. At the same time, it was noted that they do share some properties. Thus the question how such constructions are related to each other in the overall ‘fabric’ of the grammar becomes an urgent one. Trying to derive them by means of general rules is simply out of the question in view of the facts, but just listing the constructions as ‘atoms’ of the grammar does not do justice to the facts either, and would not capture the generalizations that are clearly there. So we will have to find some alternative way of conceiving of relations between grammatical structures.

Let us first consider the question how badly needed such an alternative conception actually is. I argued in section 3.1 that the very specific pattern *zich een weg banen door X* (with both the verb and the path-marker lexically specified), as well as the somewhat more general superordinate schema *zich een weg +V + OBL* had to be conceived of as stored in long term memory of speakers of Dutch, with the specific schema as the prototype of the general one (cf. Figure 1). At this level of specificity, the grammars of Dutch and of English appear to be organized similarly: in English, the pattern *make one’s way through X* is the prototype of *V one’s way OBL*.¹³ But already at a slightly more abstract level important differences emerge. The English *way* construction has been characterized as a specific case of resultative constructions (of the type *He cried his eyes red*, so-called fake-object resultatives; cf. Goldberg 1996:50, and references cited there, for discussion). In any case, it exhibits a transitive pattern, with two argument positions (subject and object). But the Dutch

¹³ One difference being that *make* in the English pattern is also connected to the general activity-verb *make*, inheriting any relevant properties.

way construction, with its characteristic reflexive element *zich*, exhibits a *ditransitive* pattern, with three arguments: subject, direct object and indirect object; it actually looks like a kind of benefactive construction. So for Dutch, the partial taxonomic network consisting of the prototype *zich een weg banen door X* and its superordinate *zich een weg +V + OBL* must in turn be considered subordinate to a more general ditransitive pattern, as indicated in Figure 2.

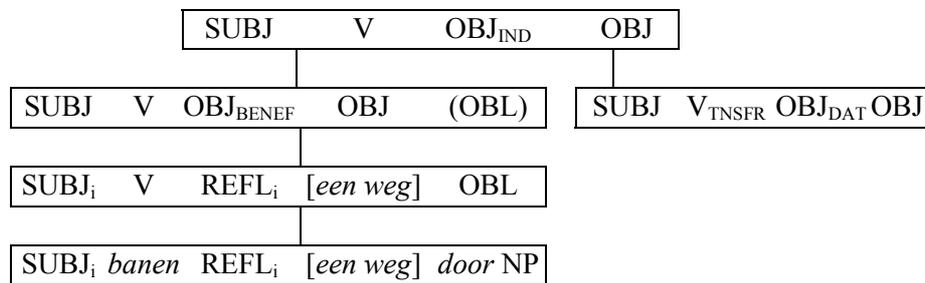


Figure 2. *The Dutch way construction as a ditransitive construction*

In English, the more general pattern to which its network of *way* constructions should be subordinated is the transitive one, as it is a kind of resultative (cf. Figure 3).

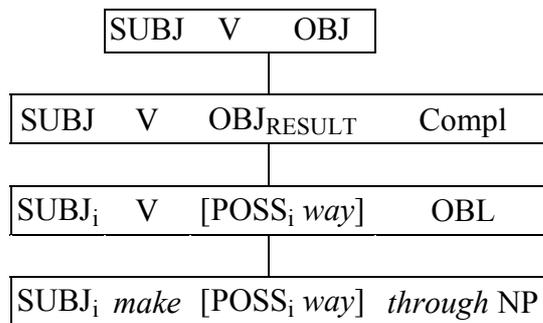


Figure 3. *The English way construction as a transitive construction*

What this suggests is that the position of these constructions in the ‘grammatical space’ of Dutch and English is quite different for each language. However, it should be noticed that the benefactive pattern near

the top in Figure 2 is not at all a productive pattern in Dutch.¹⁴ The language does have a conventional pattern *zich een weg +V*, which can be glossed as ‘to V oneself a way’ and which (roughly) means ‘create a path/opportunity for oneself (and use it), by means of V-ing’, but not the more general pattern *iemand +Y +V*, to be glossed as ‘to V someone Y’ and meaning ‘to make Y for someone by V-ing’. Curiously enough, English does have a productive pattern of this sort; while (40) is unacceptable in Dutch (this can only be expressed as in (41)), the English parallel (42) is perfectly acceptable.

- (40) ?? Jan maakte haar een boterham.
 (41) *Jan maakte een boterham voor haar.*
 John made a sandwich for her
 ‘John made her a sandwich.’
 (42) John made her a sandwich.

The paradox is that English has a rather productive general benefactive construction, but the *way* construction is not an instance of it, while Dutch does *not* have a productive benefactive construction, although its *way* construction does seem to instantiate it. In any case, the consequence is that we have to exclude the *way* construction from the network of Dutch ditransitive constructions, and replace Figure 2 by Figure 4: the network of more and less specific *way* constructions actually constitutes a kind of island in the whole of the grammar.

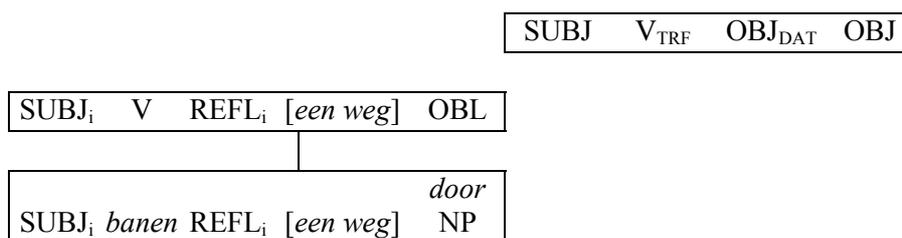


Figure 4. *The Dutch way construction island*

¹⁴ This applies to the standard language. Ditransitive patterns (with more or less specific semantics) exhibit different degrees of productivity in different regions (cf. van Bree 1981).

However, this once again produces the problem that we can no longer express the similarities that do exist between *way* constructions and ditransitive clauses. So the issue of an alternative way of conceiving of such relations is really quite urgent. What I want to suggest, at least as a tentative proposal, is to mark similarities between parts of constructions in a manner that can in no way be confused with the categorizing relationships between entire constructions in a taxonomic network. We have already seen in section 3.2 that it would be wrong to posit a template generalizing over the first and second type of *way* construction in Dutch. The point is that while the top node in the network of Figure 1 is itself a category that can serve as a template licensing new utterances, the top nodes in the hypothetical network of Figure 5 does not play such a role in the language (as shown in sections 3.2 and 3.3), and thus this kind of representation is at least misleading. Instead, I propose to indicate similarity just by means of links between the elements of the pattern participating in the relationship, as in Figure 6 – which has the additional advantage of making the internal structure of the constructions more explicit.

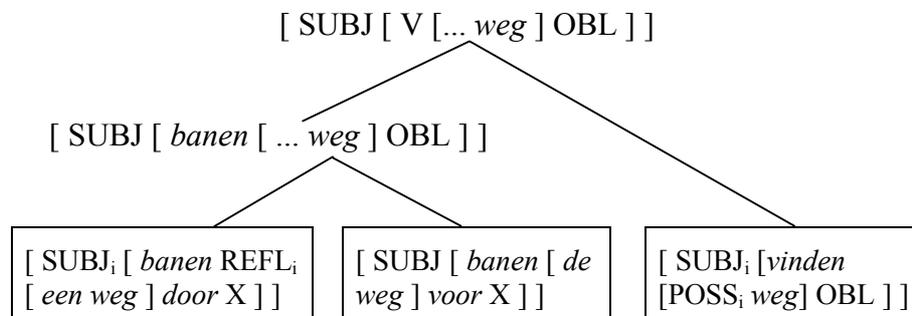


Figure 5. *Similarity as a categorizing relationship*

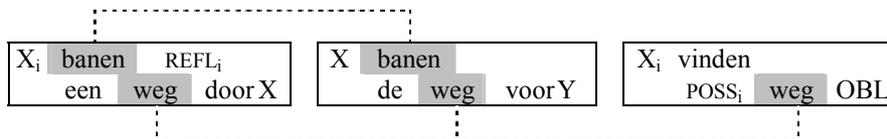


Figure 6. *Similarity as a link between parts*

This idea may be considered reminiscent of the ‘lexical redundancy rules’ in Jackendoff (1975), which were also meant to capture (morphological) relationships between words without deriving one kind of word from another (by transformational processes). However, at least one very important difference is that at the time, this proposal was believed to help keeping syntax separate from lexicon and morphology, thus ‘rescuing’ a modular view of linguistic knowledge, while we now see that this type of phenomena occurs in syntax as much as anywhere else (as noted “in passing” by Langacker 2000:20).

The considerations leading to this idea can also be seen as another instance of (a generalized form of) the argument put forward by Goldberg (2002). Goldberg argues against ‘overplaying’ generalizations over distinct surface forms that are to some extent paraphrases of each other (‘alternants’ such as *give X to Y/give Y X*, or *load X with Y/load Y onto X*), and especially against the idea (of which transformations are only one extreme implementation) that such a similarity requires a structural account in the grammar, an analysis of partly similar sets of clauses as instantiations of a single more general pattern. Goldberg’s point is that such generalizations are actually not at all as broad as they may seem at first sight, and that they moreover prevent the formulation of other generalizations that in fact hold more widely. As Goldberg recognizes, abolishing a structural account of course does not annul “the question of how the overlap in meaning between alternants is accounted for”, but the answer that does justice to the facts without overgeneralizing refers to just the shared elements themselves: “The shared meaning can be attributed directly to the shared verb involved” (Goldberg 2002:343).

Finally, this proposal is also very much in the spirit of the one made for morphology by Bybee (1995, among others), who furthermore links her ideas intimately to usage, in particular frequency (see also Bybee & Hopper 2001): the relations between similar parts of linguistic elements and their strength are, at least to a considerable extent, determined by the number of stored elements, and ultimately usage events, participating in them.¹⁵

¹⁵ It cannot be excluded a priori that the technique used by Langacker (1988, 2000, among others) of graphically marking different degrees of entrenchment of categories and degrees of strength of relations in a purely taxonomic network

Exactly the same mechanism can be used, of course, for representing similarities between structural features of distinct constructions. In Verhagen (2002:422), I gave the representation in Figure 7 for a combination of taxonomic and similarity relations, both lexical and structural, for a subset of the constructions discussed here.

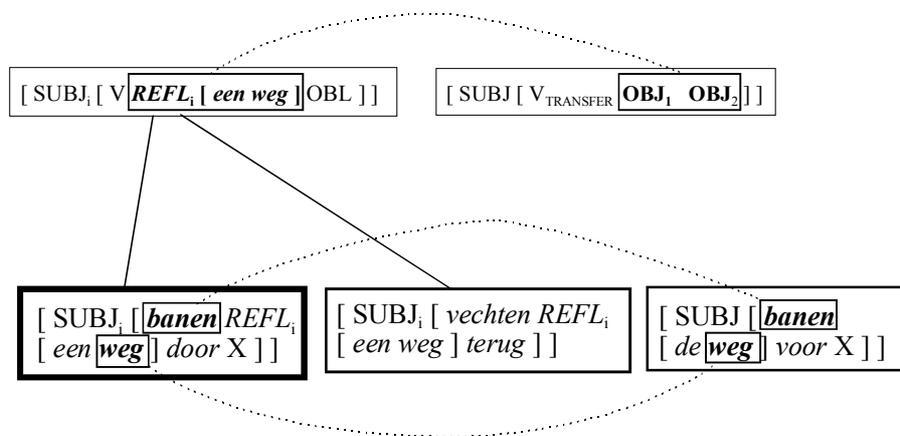


Figure 7. Network with some categorizing and similarity links

Even though it does not by far represent all relationships that the elements of the constructions participate in (representing more relations very quickly leads to obscurity of the two-dimensional picture), it suffices to show how constructions may definitely be islands in the grammar of the language, and nevertheless an integral part of the ‘fabric’ of grammar (the similarities providing bridges, so to speak). Even if similarities do not have the status of a rule telling the speaker how to *build*

will ultimately be able to do the same work, so that the distinction between a productive template and a bundle of links between similar elements can be reduced to the mechanism of differential entrenchment. Moreover, schema’s that do have the status of productive rules are, of course, also based on similarities between specific cases, and speakers may vary in what for them are only similarities, and what are conventional productive schema’s. Nevertheless I find it useful, at least for the moment, in view of the considerations presented above, to mark the distinction quite explicitly.

structures in the language, they do contribute to the overall coherence of grammar. Similarities between different stored patterns, especially if they are both formal and semantic (i.e. symbolic), increase the strength (entrenchment) of the memory representation of the patterns involved. Thus they provide constraints on the patterns allowed into the grammar: those that look more like others are more easily ‘admitted’ (as they inherit part of the required degree of entrenchment from memory structures that are already available) than those that have less resemblance to other constructions. On the other hand, the inclusion of non-standard constructions is never prohibited in any absolute sense; constructions with properties they share with others are favored, but by exactly the same token idiosyncratic ones ‘only’ require more specific direct support from actual usage. There is no reason to suppose that different cognitive mechanisms have to be involved in the development and maintenance of general and specific parts of a grammar. Different grammatical constructions in a language do not have completely random overall structures, but the variation is definitely larger than what one might expect on a purely rule-based account; in order to do justice to the facts, the latter requires distinctions (e.g. core/periphery, rule/exception) that actually have no other motivation than preserving the centrality of the role of abstract categories and rules in explanatory accounts of grammatical structure. A usage-based approach is very well capable of capturing regularities where this is appropriate; the point is that it is also very well capable, without additional mechanisms, of avoiding them where they are not appropriate – and they often are not.

Acknowledgements

I am grateful to the Netherlands Institute for Advanced Study (NIAS) for providing me with the opportunity, as a Fellow-in-Residence, to complete the work for this paper.

References

- Barlow, M. & S. Kemmer (eds.) (2000). *Usage-Based Models of Language*. CSLI Publications, Stanford, CA.
- Bree, C. van (1981). *Hebben-constructies en datiefconstructies binnen het Nederlandse taalgebied: een taalgeografisch onderzoek*. PhD Dissertation, Leiden University. Intercontinental Graphics, Hendrik Ido Ambacht.

- Bybee, J.L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes* 10, 425-455.
- Bybee, J. & P. Hopper (eds.) (2001). *Frequency and the emergence of linguistic structure*. John Benjamins, Amsterdam and Philadelphia.
- Goldberg, A.E. (1996). Making one's way through the data. In M. Shibatani & S.A. Thompson (eds.), *Grammatical Constructions. Their Form and Meaning*, 29-53. Oxford University Press, Oxford.
- Goldberg, A.E. (2002). Surface generalizations: An alternative to alternations. *Cognitive Linguistics* 13, 327-356.
- Israel, M. (1996). The way constructions grow. In A.E. Goldberg (ed.), *Conceptual Structure, Discourse and Language*, 217-230. CSLI Publications, Stanford, CA.
- Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language* 51, 639-671.
- Jackendoff, R. (1990). *Semantic Structures*. MIT Press, Cambridge, MA.
- Kramer, A. (2002). De geschiedenis van de Nederlandse *weg*-constructie. Unpublished MA-thesis, Dept. of Dutch, University of Leiden.
- Langacker, R.W. (1988). A usage-based model. In B. Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, 127-161. John Benjamins, Amsterdam and Philadelphia.
- Langacker, R.W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (eds.), *Usage-Based Models of Language*, 1-63. CSLI Publications, Stanford, CA.
- Loewenthal, J. (this volume). Meaning and use of causeless causative constructions with *laten* in Dutch. This volume, 97-129.
- Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science* 12, 49-100.
- Verhagen, A. (2002). From parts to wholes and back again. *Cognitive Linguistics* 13, 403-439.
- Verhagen, A. & S. Kemmer (1997). Interaction and causation: Causative constructions in modern standard Dutch. *Journal of Pragmatics* 27, 61-82.

On the Interaction of the Dutch Pragmatic Particles *hoor* and *hè* with the Imperative and Infinitivus Pro Imperativo

Robert S. Kirsner (University of California at Los Angeles)

1. Introduction

Although closely related to Modern English, Modern Dutch exhibits many lexical and grammatical phenomena having no counterpart in English at all. Furthermore, even though reference grammars such as the two-volume *Algemene Nederlandse Spraakkunst* or *ANS* (Haeseryn et. al. 1997) mention many of these phenomena in passing, they do not describe or analyze them in any detail. Both the linguist wishing to study how such words and structures are used and the foreign student of Dutch wanting to sound less foreign will typically find in the standard reference grammars only a vast desert.

The purpose of the present paper is to discuss two such phenomena and to explain the hitherto undescribed interaction between them. First we consider the utterance-final pragmatic particles *hoor* (literally ‘hear’) and *hè* (‘isn’t it’): cf. Kirsner & Deen (1990), Kirsner & van Heuven (1997). Next we survey two imperative structures of Dutch: the bare verb stem (STM), used with finite clause word-order, and the ‘infinitive used as imperative’ (infinitivus pro imperativo), or IPI, primarily used with verb-final word order; cf. Kirsner, van Heuven & Caspers (1998), van Heuven & Kirsner (1999). Because English does not have an equivalent of either the *hoor-hè* contrast or the STM-IPI contrast, we provide illustrative examples.

We shall then attempt to predict on semantic and pragmatic grounds how the particles will be used together with the imperatives. Which imperative structure will favor or disfavor the use of which final particles? Our predictions will be tested empirically against (i) the results of a questionnaire experiment and (ii) counts of actually occurring imperative plus particle combinations in various Dutch texts. We shall discover that, in addition to allowing us to test our hypothesis, both the questionnaire data and the text count data raise further questions not

originally considered. We will suggest possible answers to these questions. We will conclude by examining theoretical implications of the study.

2. The utterance-final pragmatic particles *hoor* and *hè*

2.1 Observations

Consider the following examples:

- (1) a. Jij komt morgen ook.
‘You are coming tomorrow too’.
- b. Jij komt morgen ook, hoor.
‘You are coming tomorrow too, mind you./
You be sure to come tomorrow also.’
- c. Jij komt morgen ook, hè?
‘You are coming tomorrow too, aren’t you?’
- (2) a. Stikstof is een gas.
‘Nitrogen is a gas.’ [Can be an “encyclopedia sentence”]
- b. Stikstof is een gas, hoor.
‘Nitrogen is a gas, mind you/ son.’
- c. Stikstof is een gas, hè?
‘Nitrogen is a gas, isn’t it?’
- (3) a. *Stikstof is een gas, hè hoor.
‘*Nitrogen is a gas, isn’t it, mind you.’
- b. *Stikstof is een gas, hoor hè?
‘*Nitrogen is a gas, mind you, isn’t it?’
- c. *Stikstof is een gas, hoor?
‘*Nitrogen is a gas, mind you?’
- (4) a. Dag. ‘Hello/Goodbye’
- b. Dag hoor. ‘Goodbye/Goodbye then/*Hello then’
- c. Dag hè? ‘Goodbye, OK? = I’m leaving now, alright?’

Examples (2b,c) demonstrate that *hè* and *hoor* are used to interact with the Hearer and are not simple statements of fact such as (2a). Consider also that the *ANS* states (Haeseryn et. al. 1998:582) that *hè* and *hoor* are

“especially used when there is a certain intimacy between speaker and hearer which makes informal language use possible”. The sentences in (3) show that *hè* and *hoor* do not co-occur and that sentences containing *hoor* cannot be questions. The sentences in (4) show that when *hè* and *hoor* are added to the greeting *dag*, the utterance must be interpreted as a final rather than an initial greeting. We now attempt to account for these facts.

2.2 Analysis of the *hè-hoor* opposition

Inspired by Schiffrin’s (1987) treatment of pairs of English discourse markers such as *now*, *then*; *I mean*, *y’know*, we may propose a maximally schematic analysis of *hè* and *hoor* (in the sense of Langacker 1991:265; 2000:4) in which they share certain characteristics but contrast on at least one level; cf. Kirsner & Deen (1990), Kirsner, van Heuven & van Bezooijen (1994). Observe again that whereas *dag* by itself can be used to communicate either an initial greeting (*Dag meneer Janssen!* ‘Hello Mr. Janssen!’) or a farewell (*Dag* ‘Goodbye’), *dag* with *hoor* or *hè* can only communicate a final greeting.¹ Hence, both *hoor* and *hè* claim that contact *has already been made between Speaker and Hearer*, that there is a Speaker-Hearer relationship. And because *hoor* and *hè* do not normally combine (at least not in utterances directed towards a single hearer in a single speech act) they must contradict each other at some level of the analysis. Our analysis is given in Table 1.

¹ Note as further support for this line of argument that the dedicated *initial* greeting *hoi!* ‘hi, hello’ (Cook 1995:94) does *not* co-occur with *hoor* or *hè*: **Hoi, hoor!*, **Hoi, hè?* On the other hand, one of my consultants has indicated that the dedicated informal *final* greeting *doei* can be used with *hoor* at least: *Nou doei, hoor* ‘Well, toodle-oooh/bye-bye, then’. (The apparent oddness of *?Nou doei, hè?* – rejected by my consultant – suggests that there is something non-negotiable about the finality of *doei*, that the Speaker need not seek confirmation from the Hearer.)

Contact has already been made between speaker and hearer	
Speaker focuses hearer's attention on material prior to particle	
Speaker explicitly <i>asks</i> hearer for something (=confirmation or acknowledgement)	Speaker explicitly <i>does not ask</i> hearer for anything
___ <i>hè</i>	___ <i>hoor</i>

Table 1. *Schematic semantic analysis of the pragmatic particles hoor and hè*

Two final aspects of *hè* and *hoor* deserve mention. First, a complete understanding of these particles is only possible when one considers the full range of final particles available to Dutch speakers, such as *zeg* 'say' and *joh* 'buster'; cf. Kirsner & van Heuven (1996, 1999). Second, the diachronic evolution of *hoor* seems to be much clearer than that of *hè*. According to the *Woordenboek der Nederlandsche Taal* (Beets & Knuttel 1912:1086), *hoor* evolved from *hoort ge* 'you hear, do you hear' much in the way that English *goodbye* evolved from *God be with ye*. No clear etymology is given for *hè*, even though some native speakers opine that it comes from *heus* 'really.' In any case, the shift of *hoor* from literal 'hearing' to 'heeding' or 'paying attention' is quite reasonable: cf. Sweetser (1990:34-35). Because the final particle is entirely optional, explicit use of it can suggest (by Gricean maxims) that the Speaker is explicitly calling the Hearer's attention to something which the Hearer seems to be unaware of.

2.3 Subuses of *hoor*

The particle *hè* functions to some degree like an English tag-question and might therefore be relatively easy to understand. But because English has nothing whatsoever like *hoor*, it may be useful to briefly illustrate various conventionalized exploitations (cf. Buitenhuis 1993):

- (5) URGING.
Zegt u het maar (, hoor) [counterperson in a sandwich shop]
say you it but (, hear) = ‘What will it be?’ (Go ahead and (do) tell
me what you want to order!) Can also be FRIENDLINESS; see also
SIMPLE EMPHASIS below.
- (6) CORRECTION (which can either be unfriendly or friendly, de-
pending on the exact nature of the personal relationship between
Speaker and Hearer. It may be implied that the Hearer should have
been aware of what the Speaker is telling him.)
Stikstof is een gas, hoor!
‘Nitrogen is a gas, son.’
Note that the correction here is of an unspoken assumption rather
than an explicit statement. Cf. stressed *wel* to counter the explicit
negators *niet* or *geen*, as in (7):
- (7) Stikstof is wel een gas!
‘Nitrogen is too a gas!’ (You said it wasn’t).
- (8) REASSURANCE
Je krijgt van de tandarts wel een verdoving, hoor
(from Buitenhuis 1993)
you get from the dentist indeed an anesthetic, hear.
‘The dentist will surely give you an anesthetic, son/my dear.’
- Part of the message of reassurance comes from the explicit profiling of a
personal relationship between Speaker and Hearer. *Hoor* functions as
“linguistic touching”, a kind of linguistic pat on the shoulder or arm.
- (9) SIMPLE EMPHASIS THROUGH INTERACTION WITH HEARER
Ma heeft de TV voor twee weken ingepikt. We moeten lezen!
Onze taal beviel haar niet en de TV krijgt de schuld. Ze zegt dat
we er stom van worden! Oenig hoor!
[From comic strip *Door dik en dun*.]
‘Mom has taken away the TV for two weeks. We have to read!
She didn’t like the way we were talking and she blames TV. She
says that it makes us stupid. Dumb hear!’

- (10) WARNING
- a. Dit gaat fout (hoor)
 this goes wrong = ‘This isn’t working right’
 (, so do something!)
- b. Die beker valt om (hoor)
 that cup falls over (hear) = ‘That cup’s spilling’
 (, so watch out/ grab it, etc.)’

3. The verb stem imperative (STM) versus the ‘infinitivus pro imperativo’ (IPI)

There are numerous ways of communicating commands and requests in Dutch; cf. (11a,c-e). In this paper, we shall consider only structures (d) and (e):

- (11) a. Verb-first: Loopt u door!
 ‘walk you through’ = ‘Walk to the rear of the bus!’
- b. Verb-first: Loopt u door?
 ‘Are you walking to the rear?’ [With final rise]
- c. Verb-second: U loopt door!
 ‘You are walking through! You are walking to the rear of the bus!’
- d. STM: Loop door!
 ‘Walk through’
- e. IPI: Doorlopen
 ‘(to) through-walk’ (IPI ‘Infinitivus pro imperativo’)

3.1 Observations

We may begin by noting that STM has a wider range of uses than IPI; cf. the discussion in Paardekooper (1951), Proeme (1984), Duinhoven (1984), and Blom (1987). Whereas STM can be used to communicate commands and requests, conditional messages, and curses, IPI is limited to commands:

- (12) a. Hang de was buiten. ‘Hang the laundry outside.’
 (Proeme 1984) COMMANDS/REQUESTS

- b. Hang de was buiten en het gaat regenen. CONDITIONAL
‘Hang the laundry outside and it starts raining.’
(‘IF you hang the laundry outside, what happens? It starts raining.’)
- c. Vul de bon in en win een reis! CONDITIONAL
‘Fill in the coupon and win a trip!’ (de Haan 1992)
- d. Val dood! ‘Drop dead, F* you’ CURSE
- e. Krijg de tering! ‘Catch tuberculosis’ = ‘F*** off!’ CURSE

- (13) a. De was buiten hangen. COMMAND
b. *De was buiten hangen en het gaat regenen.
c. *Doodvallen!/ *De tering krijgen!

The imperative construction with STM is characterized as being addressed to some Hearer, some specific person with some specific time period implied. (Note that the word order is that of a finite – i.e. tensed – clause.) With the IPI, no one in particular is (felt to be) addressed and no specific time period is implied. Consider de Haan’s example (1992:101).

- (14) a. Houd de deur vrij. ‘Keep the door open’
keep the door free Cf. Jan houdt de deur vrij.
‘John keeps the door open.’
b. De deur vrij houden. ‘Keep the door open’
the door free to keep

Blom (1987:185) characterizes *Rook niet!* ‘Smoke not’ as the personal advice of a proselytizing ex-smoker urging his interlocutor to change his behavior, while *Niet roken!* ‘No smoking’ is a public announcement meant to regulate behavior in a public place. Paardekooper (1951:100-101) states that whereas *Jongens, kom binnen en maak je huiswerk* ‘Boys, come in and do your homework (STM)’ could be said by a parent directly to his or her children, *Jongens, binnenkomen en je huiswerk maken* (IPI) is an indirect command, perhaps a parent’s command being repeated by and relayed by an older sibling.

An important element of IPI is the sense that the action is part of some standard procedure of some sort, hence capable of being repeated. Blom (1987:182) contrasts *Schenk jezelf een borrel* in ‘Pour yourself a drink’, which might be said by a sincere host or hostess to a guest at a dinner party, with *Jezelf een borrel inschenken*, which in turn could be

said by a director of a play to remind an actor that the character the actor is playing is supposed to pour himself a drink at that point in the play.

3.2 Towards an analysis of the STM-IPI opposition

We may summarize the above discussion of differences in usage between STM and IPI as in Table 2, adapted from van Heuven & Kirsner (1999: 88).

	STM	IPI
1. Relation to hearer	Personal/direct	Impersonal/indirect [i.e. absence – but not denial – of personal perspective]
2. Characterization of event	To be imagined	Explicitly an action to be undertaken
3. Status of activity	[Unspecified]	Part of some sort of standard procedure

Table 2. *Comparison of message components associated with STM and IPI*

Now one might further be able to derive these contrasts in message fractions from a more abstract, underlying contrast in meaning between the two forms, as we did for *hoor* and *hè* in the previous section.² As a first approximation, one could propose the relatively abstract, schematic semantic analysis of the STM-IPI opposition shown in Table 3 below, where the meaning for STM is adapted from Proeme (1984).

² The theoretical issue raised by such highly schematic analyses will be discussed below in section 7.1 below.

STM	IPI
Hearer must imagine self as being the source of the action or locus of the state (named by the verb)	Standard name of action or state

Table 3. Possible schematic analysis of the STM-IPI opposition

To support such an abstract analysis, one would then have to argue and demonstrate that the various concrete message components listed in Table 2 result from Gricean inferential mechanisms. Consider, for example, the fact that in (13a) versus (12a), the IPI communicates not only that the hearer should *imagine* himself or herself as carrying out the action (as in conditionals) but also constitutes a command to *actually undertake* it. This could be considered a consequence of the fact that, in choosing to use IPI the Speaker also chooses to *not* use STM, in which there is an explicit component IMAGINE, allowing for a conditional interpretation. Similarly, the direct, personal flavor of STM could be derived from the fact that STM explicitly addresses a Hearer but IPI does not. And the fact that IPI but not STM suggests a standard procedure could be derived from the fact that the bare infinitive is the standard name of an action. The infinitive, after all, is used as the citation form in Dutch dictionaries and it is the verb-form which Dutch-speaking children learn first.

But it might also be the case that the IPI is a specific construction which contrasts with other constructions and which, as a consequence, could have more semantic content than merely ‘standard name of an action’. In order to say more, one would need to survey the many other uses of infinitives in Dutch (cf. Lambooy 1962, IJbema 2002: 181-184) and see whether it is useful to propose a unitary analysis of infinitive morphology in them or not and what the consequences would be for the analysis of IPI given above. For the immediate purpose of predicting the possible interaction of STM and IPI with *hè* and *hoor*, we shall rely on the ‘lower-level’ uses given in Table 2.

3.3 Contrasts in gruffness and suddenness

Preliminary though it may be, a number of facts of usage are explained by the scheme of message components given in Table 2. For example, commands may be considered a relatively face-threatening speech act. One way of defusing the possible threat, and of being more polite, is to use non-personal rather than personal constructions; cf. Brown & Levinson (1987:191). Some native Dutch speakers judge the instruction *Doorlezen!* 'Keep on reading' (IPI), said to a pupil or student, to be less gruff, less insistent than *Lees door!* (STM). Similarly, *Nu lekker slapen!* 'Now sleep well' and *Opstaan* 'Get up', said to a child in the IPI form, suggesting a procedure rather than a single isolated action, are felt to be less abrupt, requiring less instantaneous compliance, more allowing of the unfolding of a temporal process than the STM equivalents *Slaap nu lekker* and *Sta op*. Accordingly, military commands, to be obeyed instantaneously, are typically given in the STM form rather than the IPI. One has *Geef acht!* 'Give attention' 'Attention, ten-HUT!' and *Presenteer geweer!* 'Present arms!', not *Acht geven!* or *Geweer presenteren!* A particularly interesting minimal pair is *Stop! De brug is weg!* 'Stop! (STM) The bridge has washed away' versus *Stoppen! Er staat een stopbord.* 'Stop! (IPI) There is a stop-sign'. Greater urgency and unpredictability are associated with stopping because the bridge has suddenly disappeared than with the normal stopping-procedure one executes because there is an expected, regulation-type standard stop-sign at the intersection. Freeway signs are also instructive. Someone about to drive up a freeway off-ramp encounters as a 'wrong way' sign *GA TERUG!* 'Go back' and not *TERUGGAAN!*, but once he is safely on the highway, he may see as a 'no-passing' sign not *HAAL NIET IN!* but *NIET INHALEN*.

4. Predicting the interaction of STM and IPI with *hoor* and *hè*

We now turn to the main point of the paper, the question of which imperative structure, STM or IPI, would be more compatible with and tend to co-occur with the final particles *hè* and *hoor*. On the basis of the preceding sections, we argue as follows:

(1) With the bare STM, the activity is immediately placed in a personal perspective. The shock, as it were, is not cushioned. Hence the

instruction communicated is potentially face-threatening. Also, because the activity is not explicitly characterized as part of some standard procedure, the way it would be with IPI, it is potentially unpredicted, unexpected, surprising, hence potentially alarming, face-threatening on additional grounds.

(2) If, however, the bare STM is combined with a so-called softening particle, such as *maar* (etymologically ‘but’) or *eens* (etymologically ‘once’), the threat to face is explicitly countered.³ There is pragmatic cushioning of the shock. The hearer is urged over an abstract barrier with *maar* (cf. Foolen 1995, Janssen 1995) or told with *eens* that the action need not be repeated, hence is not as much of an imposition as it might be. Hence STM + *maar*, *eens*, etc. communicates less of a threat to face than bare STM.⁴

(3) If IPI is used, rather than STM, the activity is not put in a personal perspective, and is therefore not face-threatening to begin with. Furthermore, because the activity is characterized with IPI as a standard procedure of some kind, the action can be conceived of as somewhat predictable, hence less of a surprise, and hence even less face-threatening for a second reason. It does not ‘ambush’ the Hearer.

(4) Compared to utterances without final particles, utterances containing *hè* and *hoor* should be potentially non-face-threatening,

³ Vismans (1994) presents a valuable analysis of *maar*, *eens* and other modal particles in different kinds of directive sentences. Vismans (1993) reports on an experimental study of the relative politeness of *eens*, *ook*, *even*, *maar* and ‘no particle’ in the carrier sentences *Kun je de deur ___ dichtdoen?* ‘Can you ___ close the door?’ and *Je moet de deur ___ dichtdoen* ‘You must ___ close the door.’ Here *maar* was judged as the most mitigating, most polite modal particle, *eens* as the least mitigating and polite, with ‘no particle’ (as in *Je moet de deur dichtdoen*) as the least polite utterance. Shetter & van der Cruysse-van Antwerpen (2002: 149-150) gloss *maar* variously as ‘please do,’ ‘just’, and ‘go ahead and’, as in *Begin maar te eten* ‘Go ahead and start eating.’ Van der Wouden (1998:125) glosses *maar* as ‘feel free to’.

⁴ We note that brusque military commands, intended to be carried out immediately and without question, cannot be downtoned. *Geef maar acht!* ‘Go ahead and pay attention/ Feel free to pay attention!’ or *Geef eens acht!* are comparatively strange utterances requiring very special contexts; cf. Kirsner & Deen (1990:10, fn. 6).

because of the already established contact between Speaker and Hearer which these particles suggest.

(5) Accordingly, bare STM + *hè*, *hoor* should be a maximally incoherent, internally inconsistent combination. The brusqueness, the uncushioned threat to face communicated with bare STM should collide pragmatically with the intimacy and potential friendliness communicated with *hè* and *hoor*.

(6) STM + *maar*, *eens* + *hè*, *hoor* should be a more coherent combination. Since the potential threat to face suggested by STM is explicitly countered with *maar*, *eens*, *gerust*, etc., the combination should be less incompatible with the intimacy suggested by the final particles. For example, the gentle urging to perform the activity communicated with *maar*, *gerust*, *eens* should be augmented by the urging communicated with *hoor*, suggesting to the Hearer that it is alright to perform the action, even if he had not been previously aware that it was alright; cf. the (6) CORRECTION and (8) REASSURANCE, exploitations of *hoor* discussed above.

(7) IPI + *hè*, *hoor*, in turn, should also be a maximally coherent combinations. *Hoor* and *hè* would lend a personal flavor to an utterance which, by itself, is not impersonal, but non-personal, i.e. neutral and colorless. Since IPI by itself does not communicate anything that would be a threat to the Hearer's face, and *hè* and *hoor* also suggest that the Speaker-Hearer exchange is not face-threatening, the combination of these two 'hints', both suggesting a lack of threat, should not be incompatible. We summarize in Table 4, following:

	Without final particle	With <i>hè, hoor</i>
	Unmarked (there is nothing to ‘collide’ pragmatically with STM or IPI)	Potentially not Face-Threatening (because of claim of already established contact between Speaker and Hearer)
STM (1) Activity placed in a personal perspective, hence Face-Threatening. (2) Activity need not be a procedure and can be unexpected, hence Face-Threatening. (3) Is not explicitly toned-down.	Potentially Face-Threatening <i>Doe de deur dicht!</i> ‘Close the door!’	RELATIVELY INCOHERENT, INTERNALLY INCONSISTENT COMBINATION <i>*Doe de deur dicht, hoor!</i> <i>*Doe de deur dicht, hè?</i>
STM + <i>maar, eens</i> , etc.. (1) and (2) as above. (3) But Threat to Face is explicitly countered by the meaning of the optional particle, assuring the Hearer by one means or another (<i>maar, eens</i>) that the Hearer’s face is not threatened.	Potential threat to Face explicitly denied. <i>Doe de deur maar dicht!</i> ‘Feel free to close the door!’	COHERENT, SYNERGISTIC COMBINATION. (<i>Hoor, hè</i> augment effect of <i>maar, eens, gerust</i>) <i>Doe de deur maar dicht, hoor!</i> <i>Doe de deur maar dicht, hè?</i> ‘Feel free to close the door, mind you/okay?’
IPI (1) Activity not put in a personal perspective, hence not Face-Threatening. (2) Activity is a procedure, hence familiar, predictable and not Face-Threatening.	Relative absence of threat to Face. <i>De deur dichtdoen!</i> Close the door!	COHERENT COMBINATION. <i>Hoor, hè</i> lend a personal flavor to an utterance which, by itself, is absolutely colorless. <i>De deur dichtdoen, hoor!</i> ‘Remember to close the door!’ <i>De deur dichtdoen, hè?</i> ‘You’ll close the door, won’t you?’

Table 4: Predicted Interaction of STM, IPI with *hoor, hè*

5. The questionnaire experiment

5.1 Design

The first test of the hypothesis was an exploratory questionnaire administered to 78 native Dutch speakers. The design was as follows:

- 2 different test predicates (activities, lexicon): *De deur dichtdoen* ‘close the door,’ *De fiets wegzetten* ‘put the bike away’
- 2 different imperative forms, IPI or STM, as in *De deur dichtdoen*, *De fiets wegzetten* vs. *Doe de deur dicht*, *Zet de fiets weg*.
- 3 possible final particle conditions: None versus *hè* versus *hoor*.
- 2 possible modal particle conditions, namely *maar* ‘but’ feel free to’, or its absence, as in *Doe de deur maar dicht* ‘Feel free to close the door, go ahead and close the door’ or *Zet de fiets maar weg* ‘Feel free to put the bike away, go ahead and put the bike away’; *De deur maar dichtdoen* or *De fiets maar wegzetten* versus *Doe de deur dicht*, *Zet de fiets weg*, *De deur dichtdoen*, *De fiets wegzetten*.

We thus have 2 predicates x 2 grammatical forms x 3 final particle conditions x 2 modal particles = 24 conditions = 24 separate test sentences. The test sentences were listed in one random order on one version of the questionnaire and in the mirror image of that random order on a second version. Each version was administered to 39 subjects.

The subjects were asked to rate all the sentences on the following two scales:

Scale 1: Does the sentence in your opinion contain a friendly request, an authoritarian command, or something in between? Assign a ‘score’ to the sentence as follows:

VERY FRIENDLY		VERY AUTHORITARIAN
REQUEST	<== 1 2 3 4 5 6 7 8 9 ==>	COMMAND

We shall call this scale the Imperativity scale

Scale 2: How easily can you think of situations in which the sentence – exactly as written out here – would be said? Is the sentence unusual,

strange or is it common, normal? Assign a ‘score’ as follows:

UNUSUAL. STRANGE.		NORMAL. MANY CONTEXTS OR
NO CONTEXT OR	<== 1 2 3 4 5 6 7 8 9 ==>	SITUATIONS
SITUATIONS POSSIBLE		POSSIBLE

We shall call this scale the Usualness scale. Note that the relatively normal linguistic task (namely that of interpreting the sentence in some way or other) had to be carried out before the more metalinguistic task of judging whether one would ever say such a thing in Dutch and, if so, how ‘normal’ it seemed.

Note further that, in order to create a kind of baseline against which to compare the other combinations, we included in the stimulus-sentences combinations of IPI + *maar*, which were excluded from Table 4.⁵

5.2 Initial results

Let us begin by briefly considering the Imperativity judgements. The results are plotted in Figure 1.

⁵ An appreciable number of our Dutch consultants considered these combinations totally ungrammatical. The explanation for this might involve the lack of inherent temporal limitation (the non-finiteness) of the infinitive. In any case, it is suggestive that van der Wouden’s gloss for the modal particle *maar* (1998:125) does not work well in English nonfinite constructions either, presumably because it foregrounds the *feeling free* at the expense of what is to be done, so that it can no longer function as a nudge or encouragement to undertake the action: cf. *Shut the door, Feel free to shut the door, It is important to shut the door ?It is important to feel free to shut the door, Shutting the door is important, ?Feeling free to shut the door is important.*

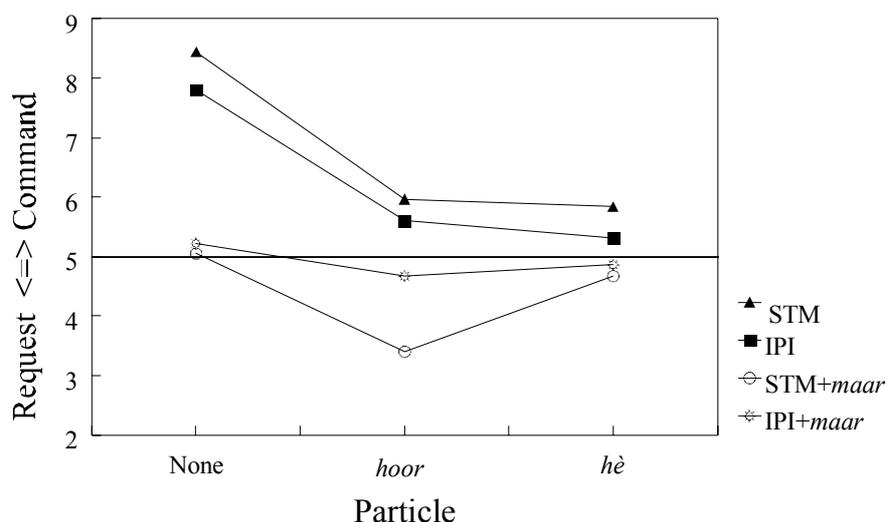


Figure 1. Mean Imperativity Scale scores for both lexical sentences combined broken down by Final Particle, Grammatical Form, and Modal Particle

For present purposes, it is enough to note that the bare IPI is judged to be between 0.4 and 0.6 of a scale point lower in Imperativity than the bare STM across all three final particle conditions (no particle, *hoor*, *hè*), and that the STM + *maar* combination is between 1.2 and 3.4 scale points lower in Imperativity than the bare STM across all three final particle conditions. That is, the decrease in Imperativity as one goes from STM to STM+*maar* is far greater than the decrease in Imperativity as one goes from STM to IPI.

Figure 2 plots the results for the Usualness judgements. Because these judgements constitute an empirical test of our analysis of the interaction of final particle with imperative form, we consider them in somewhat greater detail:

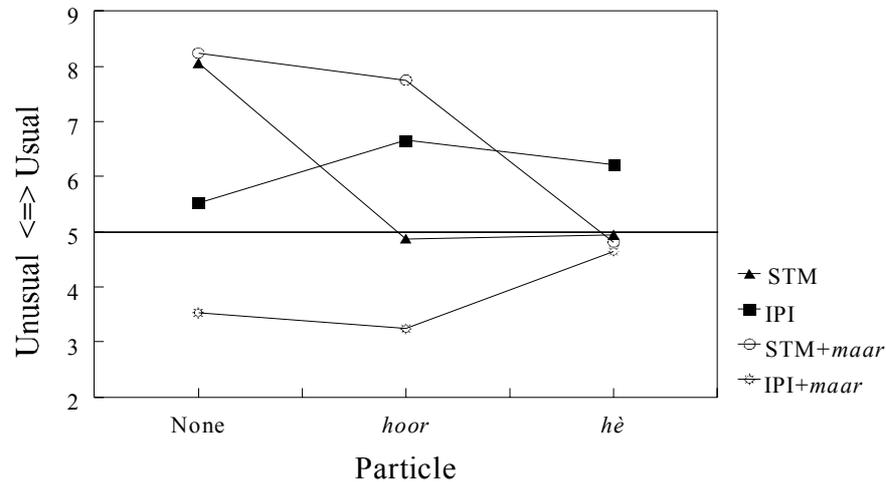


Figure 2. Mean Usualness Scale scores for both lexical sentences combined broken down by Final Particle, Grammatical Form, and Modal Particle

A repeated measures analysis of the questionnaire data reveals that Predicate, Imperative Form, and Final Particle are all significant main effects but that the only significant first-order interaction is Imperative Form * Final Particle (as we would expect). In other words, there was no significant first-order interaction involving Predicate, no significant difference between the two Predicates (closing the door, putting the bike away) in the way the form of the imperative or the final particle chosen behaved.

A series of paired comparisons was then carried out; we examine the essential prediction first. Averaging over the two Predicates, we find that when there is no final particle, the bare stem STM, with a mean score of 8.06, is judged as significantly more Usual than the infinitivus pro imperativo IPI, with a mean score of 5.52 ($p < .001$). However, when the final particle is *hoor*, there is a cross-over: it is IPI, with a mean of 6.66, which is significantly more Usual than the bare stem STM, with a mean of 4.87, rather than the reverse ($p < .001$). And when the final particle is *hè*, IPI, with a mean of 6.21, remains significantly more Usual than STM, with a mean of 4.94 ($p < .001$). These results are entirely in

line with our basic prediction. *Hoor* and *hè* are judged as more Usual with IPI than with STM.

5.3 Further results: the unpredicted lack of parallelism between *hoor* and *hè*

Let us now examine forms with the softening particle *maar*, beginning with STM. We note that, with no final particle, STM + *maar*, at 8.24, is judged as significantly more Usual than IPI and that STM + *maar* + *hoor*, at 7.75 *remains* significantly more Usual than both STM + *hoor*, at 4.87 ($p < .001$) and IPI + *hoor* at 6.66 ($p < .001$). However, with the final particle *hè*, the Usualness score for STM + *maar* decays: STM + *maar* + *hè*, at 4.81, is *not* significantly different from STM + *hè* at 4.94. We note that this decay of STM + *maar* + *hè* was not accounted for by the discussion summarized in Table 4 above, which assumed a parallelism between *hoor* and *hè*. We shall return to this point in section 5.4.2 below.

Turning to the *maar*-forms with IPI, we note that even though a number of our experimental subjects remarked that this combination was ungrammatical, there is interesting variation. Without any final particle, IPI + *maar*, with a mean Usualness value of 3.53, is indeed significantly less acceptable than both the bare STM, at 8.06, ($p < .001$) and STM + *maar* at 8.24, ($p < .001$) but not less acceptable than bare IPI at 5.52 ($p < .26$). The combination IPI + *maar* + *hoor*, with a mean Usualness score of 3.23, is indeed significantly less acceptable than each of the three other forms with *hoor*; $p < .001$ in all cases. However, when the final particle is *hè*, the Usualness score of the combination with IPI + *maar* rises by more than a scale-point, attaining a mean value of 4.64, which is not significantly less Usual than that of STM + *maar* + *hè*, with a mean of 4.81, or of STM + *hè*, with a mean of 4.94. It is significantly less Usual only than IPI + *hè*, with a mean of 6.21.

5.4 Additional discussion

5.4.1 Imperativity judgements

Our prior characterization in section 3 of STM as the most forceful imperative is confirmed by the observed relative ranking in Imperativity of STM > IPI > STM+*maar* under all conditions (bare, combined with *hoor*, and combined with *hè*). But we cannot be entirely satisfied with

this finding because the data plotted in Figure 1 go beyond this hypothesis. Note first that the reasoning summarized in Table 4 was not precise enough to enable us to predict whether STM+*maar* would be weaker or stronger in Imperativity than IPI. Second, in not saying anything about IPI+*maar*, Table 4 obviously made no prediction about the relative ranking of this combination with respect to bare IPI. And third, in lumping *hoor* and *hè* together as two pragmatic particles profiling Speaker-Hearer contact, Table 4 made no prediction that *hoor* and *hè* might behave differently from each other. Nevertheless, we suggest here that the observed data may be still explained by considering the properties of the particles and the imperative structures given in Tables 1, 2, and 3, even if all the implications of those properties were not taken into account in Table 4.

Let us consider the first point above. In retrospect, one reason STM+*maar* might be judged as less of a command than IPI is that STM contrasts with IPI in not stating explicitly that the action in question is one to be undertaken and not simply imagined (cf. Table 2). Hence, the interpretation given STM (the degree of Imperativity ascribed to it) should be *more context-dependent* than that given IPI.⁶ In the present case, the context is the addition of the polite and coaxing particle *maar*, to yield STM+*maar*. The result is that the politeness of *maar* must be interpreted together with the personal relation to the hearer communicated by STM. Since bare *maar*-less STM is still available to communicate true commands, it is reasonable that STM+*maar* is taken as contrasting maximally with it, namely as a request. IPI, however, remains impersonal and more explicitly a command, so it is not all that surprising that it is still judged as a command rather than a request.

Turning now to the second point, we might suggest that the position of IPI + *maar* exceedingly close to the midpoint of the Imperativity scale under all three final particle conditions (namely at 5.22, 4.67, and 4.86) reflects the fact that this combination is simply considered ungrammatical by many native speakers, so that they don't quite know

⁶ For a demonstration that this is case when the 'context' is the pitch level of the utterance, see van Heuven & Kirsner (1999:87-88). The degree of perceived Imperativity of STM is much more influenced by pitch level than the perceived Imperativity of IPI is.

how to judge it.⁷ (As is apparent from Figure 2, IPI+*maar* has the lowest mean acceptability of the four combinations of imperative and particle considered.)

Let us now turn to the most striking lack of parallelism in the behavior of *hoor* and *hè* seen in Figure 1, namely that between STM+*maar*+*hoor* and STM+*maar*+*hè*. The fact that STEM + *maar* + *hè* lies near the midpoint of the Imperativity scale, at 4.67, considerably above STM + *maar* + *hoor* at 3.40, might, as in the case of IPI+*maar* just discussed, reflect the relative incoherence of this specific combination of imperative form, modal particle, and final particle. We take up this issue in more detail in the following section.

5.4.2 Usualness judgements

As noted above, the data plotted in Figure 2 are consistent with our prediction that the utterance-final pragmatic particles *hoor* and *hè* would combine more felicitously with IPI than with STM. We shall now take up two questions involving *unpredicted* differences between *hoor* and *hè* and between imperative structures with and without *maar*.

The first question is: Why is the combination of STM + *maar* + *hoor* so much better (more Usual) than STM + *hoor* while both STM + *hè* and STM + *maar* + *hè* have the roughly the same, lower, Usualness score as STM + *hoor*? The answer would seem to be as follows: Because *maar* softens a command to a request (as seen in Figure 1), the use of *hoor* with STM + *maar* is less internally contradictory than with the bare STM, as was predicted in earlier discussion; cf. Table 4. According to Janssen (1995) and Foolen (1995), *maar* indicates that there is some sort of an abstract barrier or threshold to carrying out the action. Compare the *ANS* (Haeseryn et. al. 1997:457) which characterizes *Geef die boeken hier* ‘Give those books here’ as a command but *Geef die boeken maar hier* as a reassuring and friendly request. In using the STM form of the

⁷ Compare Osgood’s Semantic Differential tasks (Osgood, Suci & Tannenbaum 1957), in which subjects had to rank stimulus words (e.g. *knife*, *swamp*) on 7-point bipolar scales such as *strong-weak*, *good-bad*, *wise-foolish*, etc. If the property being measured on the scale had nothing whatsoever to do with the stimulus word being ranked, the word might receive an intermediate rank rather than one at either extreme; cf. Weinreich (1959), especially section 2, entitled ‘Is a Knife Humble or Proud?’.

imperative together with *maar*, the Speaker ‘coaxes’ the Hearer, as it were, over this abstract threshold. The use of the final particle *hoor* accordingly works *synergistically* with *maar* because *hoor* can suggest that the Hearer was unaware of something: in this case, that it is not only permitted but also desirable to cross the threshold and carry out the action in question.

Hè, in contrast, works *antagonistically* – a fact which was not sufficiently kept in mind in the discussion summarized in Table 4. On the most basic, literal, level, in using *hè* together with the more direct STM, the Speaker asks the Hearer to (i) explicitly confirm or acknowledge that he is being ordered or requested to carry out an action and (ii) to agree to it. This request for acknowledgment or confirmation *undercuts* the pragmatics of commanding. It is incoherent with a pure command, just as *hè* is incoherent with genuine (pure) questions as opposed to a statement which the Hearer is being asked to confirm or disconfirm; compare **Is stikstof een gas, hè?* ‘Is nitrogen a gas, isn’t it?’ versus *Stikstof is een gas, hè?* ‘Nitrogen is a gas, isn’t it?’ But even if *hè* is interpreted in STM + *hè* not literally but metalinguistically, on the level of speech acts (cf. Sweetser 1990:70 and passim), as in *Nou dag, hè?* ‘Well, goodbye, OK?’ where the Speaker asks the Hearer to acknowledge and assent to his Speech Act of saying goodbye, there is something decidedly strange about the Speaker asking the Hearer to acknowledge a speech act of *commanding*.⁸ It certainly is odd with unquestionable commands, such as in a military context; cf. *Ingerukt mars!* ‘Dis-MISSED!’ versus *???Ingerukt mars, hè?* ‘Dis-MISSED, okay?’ Hence, it makes sense that STM + *hè* should be judged as roughly on the same (low) level of Usualness as STM + *hoor*.

But *maar* even adds to the incoherence. If *maar* gently ‘coaxes the Hearer over a barrier’ to performing the action, lowering the Imperativity, the use it with of *hè* – asking for instant confirmation – increases the Speaker’s insistence, which makes the command *more* of a

⁸ For the sake of completeness we note that whereas the use of *hè* to simply request *confirmation* of a statement can be paraphrased with the tag *is het niet?* ‘isn’t it’, the metalinguistic use of *hè* to request *acknowledgment* cannot be. One has both *Stikstof is een gas, hè?* and *Stikstof is een gas, is het niet?* ‘Nitrogen is a gas, isn’t it?’, but alongside *Nou dag, hè?* ‘Well goodbye, alright?’ one does not say *??Nou dag, is het niet?* ‘Well goodbye, isn’t it?’ cf. Kirsner (2001:21, fn.5).

command, thereby leading to a pragmatic collision. If we follow van der Wouden (1998:125) and translate *maar* with English *feel free to*, we observe that, in contrast to *Close the door!*, *Close the door, will you?*, and *Feel free to close the door*, the combination *Feel free to close the door, will you?* is exceedingly strange. The kind invitation of *feel free to* (the rough English analogue of *maar*) is undercut by the nagging, insistent character of the *hè*-like tag-question. Presumably the same sort of mechanism underlies the internal inconsistency of *Doe de deur maar dicht, hè?*, *Zet de fiets maar weg, hè?*

A second question suggested by scrutiny of Figure 2 is: Why – if many subjects consider IPI+*maar* to be ungrammatical – does IPI + *maar* + *hè* nonetheless score better in Usualness than IPI + *maar* or IPI + *maar* + *hoor*? Here one might cautiously suggest that the reason is that (i) unlike STM, IPI is less personal, hence potentially less direct, and (ii) that it is here combined with two softeners, *hè* and *maar*, both of which operate to convert the name of a procedure, closing the door, into a suggestion for action rather than a command.

6. Quantitative data from texts

We now return to our main theme of the relative coherence of *hoor* and *hè* with STM and IPI. The second kind of evidence in favor of our hypothesis that *hè* and *hoor* will be more coherent with IPI than with STM is provided by text counts of the relative occurrence of STM and IPI with and without final *hoor* and *hè*.

6.1 The prediction

We predict that the more coherent combinations, IPI + *hoor*, IPI + *hè* will be more frequent in running texts than the less coherent combinations of bare STM + *hoor*, STM + *hè*. Table 5, below, presents the data for five separate corpora, covering 895 pages of text. Note that the notation STM (+ *maar*) in the table covers instances of both bare STM (without any modal particle) and STM + *maar*, *gerust*, etc. IPI here

refers, as always, to instances of IPI without *maar*: Although we encountered no final *hè* in our corpora, there were 8 instances of *hoor*.⁹

(1) Carmiggelt (1975) *Slenteren*

	Plain	With final <i>hoor</i>	Total	% <i>hoor</i>
STM (+ <i>maar</i>)	34	1	35	3
IPI	8	2	10	20

(2) Hellinger (1967) *Vlammen*

	Plain	With final <i>hoor</i>	Total	% <i>hoor</i>
STM (+ <i>maar</i>)	40	0	40	0
IPI	4	1	5	20

(3) van Straten (1989) *Lukt het Agnes?*

	Plain	With final <i>hoor</i>	Total	% <i>hoor</i>
STM (+ <i>maar</i>)	62	0	62	0
IPI	27	1	28	4

(4) Campert (1960) *Een ellendige nietsnut en andere verhalen*

	Plain	With final <i>hoor</i>	Total	% <i>hoor</i>
STM (+ <i>maar</i>)	21	0	21	0
IPI	4	1	5	20

(5) Reve (1972) *De Avonden. Een winterverhaal*

	Plain	With final <i>hoor</i>	Total	% <i>hoor</i>
STM (+ <i>maar</i>)	161	1	162	1
IPI	26	1	27	4

Table 5. *The distribution of STM and IPI with and without hoor in 5 Dutch texts*

We observe that the skewing is the same in each text: the percentage of IPI with *hoor* is higher than the percentage of STM (+*maar*) with *hoor*. The possibility that this distribution could occur by chance is analogous to the possibility of flipping a coin five times and getting five heads in a row, namely $(2)^5 = 1/32 = .031$, which is less than .05, the customary

⁹ A further search, in an additional text, yielded one example of *hè*. See section 6.2.3 below.

threshold for statistical significance. The figures for the aggregate sample are shown in Table 6:

	Plain	With final <i>hoor</i>	Total	% <i>hoor</i>
STM (+ <i>maar</i>)	318	2	320	1
IPI	69	6	75	8

Table 6. *The distribution of STM and IPI with and without hoor in the aggregate sample*

It should be pointed out here that our original prediction is confirmed even better than these figures indicate, for the two cases out of 320 where STM (+ *maar*) does co-occur with *hoor* contain not bare STM but only STM + softening particle. The first case is sentence (15), with the particle *gerust* ‘calmly, go ahead’, from a short story collection by Simon Carmiggelt. The second is from Gerard Reve’s classic novel *De Avonden* and contains *maar*:

- (15) Wel, wel, wel dat doet me genoeg. Wil je thee jongen? De koekjes staan op het dressoir. Neem er gerust een paar, hoor. (Carmiggelt 1975:54)

‘Well, well, well. That makes me glad. Do you want tea, my boy. The cookies are on the sideboard. Go ahead and have a couple, hear.’

- (16) Hij liet het oor los, aaide over de kop en zei iets luider: ‘Huil maar niet, hoor. Zo ver is het nog niet...’ (Reve 1972:212)
He let go of the ear, stroked its head and said more loudly: ‘There, there don’t cry. (Literally: Cry but not, hear). We haven’t got there yet...’ (Reve n.d.:184).

Accordingly, while 8% of the 75 instances of IPI contain *hoor*, 0% of the instances of *bare* STM (without modal particle) contain *hoor*. This makes sense if, as we argue, IPI + *hoor* is a less internally contradictory combination than STM + *hoor*. An example of IPI + *hoor* found in the texts is the following (Hellinger 1967:119):

- (17) ‘Oh, wat een schattig autootje. Net speelgoed. Mag ik er eens in rijden?’

Vooruit dan maar. Misschien vrolijkte het haar wat op. Ik liet haar zien hoe het monstertje schakelde.

‘Maar hier op het parkeerterrein blijven, hoor.’

‘Natuurlijk Sid.’

‘Oh, what a cute little car. Just like a toy. Might I drive it?’

Well alright. Perhaps it would cheer her up a bit. I showed her how the little demonstration car shifted gears.

‘Just stay here in the parking lot, mind you.’

‘Of course, Sid.’

6.2 The importance of negation

Just as was the case with the questionnaire data, we find that further scrutiny of text count data uncovers important trends not anticipated when our original hypothesis was formulated. We find, namely, that negative commands or prohibitions seem to favor both the occurrence of IPI versus STM and the presence of *hoor*. Consider Table 7.

Positive Commands

	STM (+ <i>maar</i>)	STM (+ <i>maar</i>) + <i>hoor</i>
Carmiggelt	32	1
Hellinger	38	0
v Straten	55	0
Campert	20	0
Reve	143	0
<i>Aggregate</i>	288	1 (0.3%)

	IPI	IPI + <i>hoor</i>
Carmiggelt	6	0
Hellinger	2	1
v Straten	16	0
Campert	2	1
Reve	20	0
<i>Aggregate</i>	46	2 (4.2%)
Prohibitions (with Negation)		
	STM (+ <i>maar</i>)	STM (+ <i>maar</i>) + <i>hoor</i>
Carmiggelt	2	0
Hellinger	2	0
v Straten	7	0
Campert	1	0
Reve	18	1
<i>Aggregate</i>	30	1 (3.2%)
	IPI	IPI + <i>hoor</i>
Carmiggelt	2	2
Hellinger	2	0
v Straten	11	1
Campert	2	0
Reve	6	1
<i>Aggregate</i>	23	4 (14.8%)

Table 7. *The distribution of STM and IPI with and without hoor broken down into positive commands and prohibitions*

As one moves from positive commands to negative prohibitions, the total percentage of IPI increases from 48/337 or 14% to 27/58 or 47% and the total percentage of *hoor* increases from 3/337 or 1% to 5/58 or 9%.

6.2.1 The favoring of IPI

One might suggest three motivations for more than three-fold increase of IPI in prohibitions. The first and most obvious factor is iconicity. Having the negator in initial position immediately clues the hearer that the

message is going to be a prohibition and prevents any confusion with a positive command. That is, one could argue that a sentence like *Niet schieten!* 'Not to shoot = Don't shoot' (with the *Niet* being the very first word the Hearer receives) is less easily confused with a positive command than *Schiet niet* 'Shoot not', in which the morpheme *schiet* could be the first word in the positive commands *Schiet!* or *Schieten!* 'Shoot.'

A second possible factor might be some sort of given-new strategy. That is STM + *niet* (with the lexical verb in initial position) might be used when the activity described by the verb has been an earlier topic of conversation and *niet* + IPI might be used when the activity described by the verb has not been. It is at least suggestive that in van Straten (1989) one can find examples such as the following. In (18) *zeg* (STM) follows an earlier instance of *zeggen* 'say' but in (19) *tobben* 'worry' has not been previously mentioned:

- (18) 'Als Johan nou weer belt, wat zeg ik dan'
 'Weet ik niet. Zeg maar helemaal niks. Dat je het niet weet.'
 (van Straten 1989:141)

'If Johan should call again, what do I say?'
 'I don't know. Don't say anything at all (STM)
 (Say) that you don't know.'

- (19) En opeens keek hij Agnes aan, met wijd opengesperde ogen en zei een beetje plechtig: 'Agnes..ik scháám me zo!'
 Braaf zei Agnes dat er voor Daniël werkelijk helemaal niets was om zich voor te schamen.
 Maar wat herkende ze die schaamte.
 'Niet meer tobben,' zei ze, 'lekker slapen.' En ze deed zijn bedlampje uit.
 (van Straten 1989:34)

And suddenly he looked at Agnes, with wide open eyes, and said somewhat solemnly: 'Agnes, I am so ashamed!'
 Decently Agnes said that there was nothing that Daniel needed to be ashamed about.
 But how she recognized that shame.
 'No more worrying = Don't worry anymore (IPI),' she said, 'to sleep tight (IPI).' And she turned off his little bedside lamp.

Finally, a third factor might be that IPI in its entirety is something of an innovation. It did not exist in Middle Dutch (Stoett 1923:241; Weijnen 1971:98). Consequently, with at least certain verbs, the older STM imperative seems old-fashioned and the IPI colloquial.¹⁰ A preacher might still say *Zondig niet meer* ‘Sin no more’ in a Bible lesson, but this phrasing is much more formal than *Niet meer zondigen* ‘No more sinning’, where, however, the use of IPI would clash with the solemnity inherent in the lexical meaning of this particular verb. Nevertheless, by bringing in a final *hoor* (another colloquial element) to underscore the personal contact between Speaker and Hearer, the Speaker can temper this solemnity further and create a playful ironic effect: *Niet meer zondigen, hoor!* ‘No more sinning, mind you!’ It would not be possible to do this in the same way with the original STM variant (e.g. *Zondig niet meer, hoor!* ‘Sin no more, mind you!’) because of (i) the original pragmatic clash already discussed between bare STM and *hoor*, and because of (ii) the huge register class between the solemn *Zondig niet meer* and *hoor*.¹¹ Such stylistic factors as the perceived novelty of a construction and its colloquial flavor might well play a role in the further propagation of that construction.

6.2.2 The favoring of *hoor*

We now turn to the second apparent trend seen in Table 5, namely the apparent favoring of *hoor* with prohibitions in general. As noted above, the overall increase as one moves from positive commands to prohibitions is from 0.9% to 8.6 % or +7.7% . If future research shows this to be a real trend, perhaps its explanation is to be sought in Duinhoven’s claim (1997:406-7) that negative imperatives are by nature weaker than positive ones and constitute more of a recommendation than

¹⁰ See further Givón’s concept of the diachronic conservatism of negation (Givón 1979:121-142). Suppose we have a verb V with a sense A and which is evolving a new sense B. If negative environments are conservative, as Givón’s discussion of the English modal verbs suggests, then STM+NEG would tend to preserve the older A sense of the verb relative to STM. The newer, B sense of the verb would tend to be favored by the innovation IPI with or without NEG.

¹¹ An analogue would be the sarcastic quotation of the solemn, Biblical Eighth Commandment followed by *hoor*: *Gij zult niet stelen, hoor* ‘Thou shalt not steal, mind you.’ – cf. Haiman’s 1990 discussion of quotation in sarcasm.

a genuine command. If this were true, a negative imperative could be potentially less of a threat to the Hearer's face than a positive command and more compatible with other elements, such as *hoor*, whose use also indicates that the Hearer's face is not being threatened.

Though the sample is perhaps too small to generalize from, there is an interesting asymmetry worth further study. The percentage of *hoor* with STM (+ *maar*) rises from $1/239 = 0.3\%$ to $1/30 = 3\%$ while the percentage of *hoor* within the IPI category rises from $2/48 = 4\%$ to $4/27 = 15\%$. This 12% difference between IPI and STM in the rate of attraction of *hoor* in prohibitions might be explained by the same mechanism outlined above. If negative imperatives involve less of a threat to face than positive ones, IPI would seem to be the form of choice and there would be a synergy between the use of IPI rather than STM(+*maar*) and the use of *hoor*. A example from Reve:

- (20) Viktor dronk met een vertrokken gezicht. 'Niet de boel verpesten, hoor,' zei Frits. 'Zo'n kwaad smoel kunnen we hier niet hebben...' (Reve 1972:117)
 Victor pulled a face as he drank. 'Now don't spoil things,' (Literally: Not the matter mess-up, hear) said Frits. 'We can't have dirty looks like that here...' (Reve n.d.: 100)

6.2.3 A note on *hè*

In making counts of the kind reported here, one discovers that final particles may be quite rare in a text, so that no data are obtained on their combination with STM and IPI. As noted already, out of 395 imperative structures counted, only 8 contained *hoor* and 0 contained *hè*. That *hè* does in fact occur with imperatives is shown by an additional count of the first 100 pages of Dorrestein (2000). This time, IPI + *hè* was encountered, but not IPI + *hoor* (and of course neither bare STM + *hè* nor bare STM + *hoor*). Though the numbers are too small to be significant, the distribution is similar to that seen for *hoor*. It is tantalizing that the sole example of IPI + *hè* which should occur is in a *negative* command, a (softened) prohibition. The data are displayed in Table 8.

Positive Commands	STM (+ <i>maar</i>)	STM (+ <i>maar</i>) + <i>hè</i>
	34	0
	IPI	IPI + <i>hè</i>
	4	0
Prohibitions	STM (+ <i>maar</i>)	STM (+ <i>maar</i>) + <i>hoor</i>
	2	0
	IPI	IPI + <i>hè</i>
	1	1

Table 8. *The distribution of STM and IPI in the first 100 pages of Dorrestein (2000) with and without final particles, broken down into positive commands and prohibitions*

The datum in question is (21):

- (21) ‘En je moeder geen verdriet doen, hè?
 And your mother no pain do, eh?
 ‘And you won’t cause your mother any distress, will you?’
 (Dorrestein 2000:38)

7. Some theoretical implications

In this paper, we have explored what could be called the ‘ecology’ of linguistic forms. After introducing the utterance-final pragmatic particles *hè* and *hoor* and after sketching the contrast in usage between the verbstem imperative and the infinitive used as an imperative, we focused on the interaction of these two components of Dutch grammar. On the basis of a semantic analysis of *hè* and *hoor* and of STM and IPI, we argued that the use of these particles would be more compatible with – and hence also more frequent with – IPI than with STM. Our prediction was borne out by the results of a questionnaire experiment with native Dutch consultants and by frequency counts of combinations of STM and IPI with and without *hè* and *hoor* in a number of Dutch texts. In addition, examination of both the questionnaire data and the text count data revealed other, unanticipated phenomena for which we sketched possible explanations. We shall conclude this paper by discussing several theoretical issues which are raised by these phenomena.

7.1 Maximalist (bottom-up) linguistics versus minimalist (top-down) linguistics

Consider first of all our prediction in Table 4 that both the combination STM + *maar* + *hoor* and the combination STM + *maar* + *hè* would be acceptable. This prediction was based on the assumption that, because both *hè* and *hoor* claim that contact has already been made between Speaker and Hearer, both particles would indicate that less of a threat to the Hearer's face was present and, hence, *both* should be compatible with *maar*, a particle used to 'coax' the Hearer over an imaginary barrier, thereby decreasing the threat to the Hearer's face posed by STM. As we discussed at some length above, this prediction was incorrect; Figure 2 shows that although sentences like *De deur dichtdoen, hoor* and *De Deur dichtdoen, hè?* are judged similarly (both ranking above sentences like *Doe de deur dicht, hoor* and *Doe de deur dicht, hè?*, which are also judged similarly on the Usualness scale), sentences like *Doe de deur maar dicht, hè?* are judged as significantly *less* Usual than *Doe de deur maar dicht, hoor*. The assumed parallelism between *hè* and *hoor* thus breaks down. The reason why STM + *maar* + *hè* is worse than STM + *maar* + *hoor* can be explained only (as we did above) by considering not what *hè* has in common with *hoor* but how they differ. This entails scrutiny of the particular meaning signaled specifically by *hè*, roughly REQUEST FOR CONFIRMATION/ACKNOWLEDGEMENT, and the particular uses which this schematic meaning gives rise to. Our initial error was thus in failing to recognize the greater importance of *lower level* schemas 'in the computation or evaluation of novel expressions' (Langacker 1991:286; cf. also Langacker 2000:29-31), with the 'novel expressions' being in this case the different imperative structures with the final particles and with and without *maar*.

This point is perhaps reinforced by examining STM + *maar* + *hoor* and IPI + *hoor*, both of which ranked – as we predicted – appreciably above STM + *hoor* on the Usualness scale. From our schematic analysis, emphasizing potential threats to the Hearer's face, there was no way to predict the *additional fact* that native speakers would judge STM + *maar* + *hoor* to be significantly *more* Usual than IPI + *hoor*. On the other hand, given one native speaker's judgement that *Doe de deur maar dicht, hoor* is more condescending than *De deur dichtdoen, hoor*, one might be able to relate this to the explicit coaxing over a barrier communicated by *maar*. The fact remains that these combinations, to greater or lesser

degrees, lead a life of their own and have – almost like lexical items – emergent properties which are not entirely predictable from those of their components; cf. Langacker’s example that a *printer* ‘is not just ‘something that prints’’ (2000:38).

7.2 Degrees of idiomaticity of collocations

The data presented in Table 7 present us with a number of new puzzles which might not be entirely solvable at the schematic level at which the data are presented. On the one hand, one can argue (as we did) that, with negation, the IPI is a handy thing to have, for it allows the negator to stand iconically in initial position, where its alerting, warning function is maximally clear. Yet before one begins a relatively abstract discussion about prohibitions and commands in general, and why the frequency of *hoor* tends to increase with negation, one might also want to know more about which *particular* verbs tend to occur with and without negation in Dutch and whether there is a tendency to favor STM or IPI in each case. How does the lexical meaning of each verb interact with the semantics of each construction and each particle? Hoeksema (1992) has argued that *all* instances of IPI, whether with negation or without negation, are ‘idiomatic,’ and are learned separately, for each verb, one by one. After reading Blom (1987), with her stimulating example of the play director, mentioned above, I am however not totally convinced that, say, alongside *Niet zeuren!* ‘No complaining!’ one could *not* – pace Hoeksema (1992:128) – also say *Zeuren* ‘Complain! You are supposed to *complain* at this point in Act 3’. Nevertheless, Hoeksema’s point is well taken. As is the case with Goldberg’s (1992) study of the English ditransitive construction, full understanding of the interaction of STM and IPI with negation and with *hè* and *hoor* would seem to require more detailed scrutiny of the Dutch verbal lexicon.

7.3 Syntax versus semantics versus pragmatics

Finally, the data of Table 7 suggest that Dutch is moving towards the situation seen in languages such as Hebrew, Spanish, and Italian (Zanutini 1997 cited in Hyams 2002) where instead of negative imperatives, one finds negator + future tense, negator + subjunctive, and/or negator + infinitive: cf. Italian *Telefona!* ‘Call!’ but *Non telefonare!* ‘Not to call’

instead of **Non telefona!* Certainly the emergence in Dutch of a fixed infinitival collocation such as the warning *Niet doen!* ‘not to do’ for ‘*Don’t!*’ (Cook 1995:142) is intriguing. Nevertheless, because Dutch *does* have productive negative imperatives such as *Ga niet weg!* ‘Don’t go away!’ alongside neg + IPI, such as *Niet weggaan!*, it seems to have escaped the attention of those formal grammarians who would explain the all-or-nothing Italian phenomena in purely syntactic terms. The skewing seen in the Dutch data presented here suggests that, whatever the synchronic end point reached, the *process towards that endpoint* is driven at least in part by pragmatic and processing considerations rather than syntax.

Nevertheless, there might also be an additional factor at work: a basic clash between negation and the pragmatic force of the imperative.¹² The presence of such a clash seems to have been considered intuitively obvious by logicians such as Hans Reichenbach, who claims (1966:342) that the symbolic formulation of the negation of an imperative is ‘meaningless.’ Compare also Han’s claim (1998:40) that ‘the directive force contributed by the imperative mood cannot be negated by a negative marker.’ Accordingly, if negative imperatives are – to at least some degree – internally inconsistent, less coherent than positive ones, and if, in consequence, less assertive forms (such as the future or subjunctive or *infinitive*) have to (or tend to) be brought into play with negation to communicate the intended negative message, one could argue that prohibitions cannot really be commanded but only suggested; cf. Duinhoven (1997) mentioned earlier.

Yet another tack would be to consider the pragmatic role that negation plays in discourse and the greater suitability of the infinitive

¹² Further evidence for the greater force of positive commands in Dutch (and hence of their greater potential threat to the Hearer’s face) is perhaps provided by special imperative constructions which cannot be negated at all. First of all, it is difficult to imagine a prohibition *Presenteer geen geweer!* ‘Do not present arms!’ alongside the military command *Presenteer geweer!* And, second, one has in Dutch the special ‘success imperative’, such as *Eet ze!* ‘Enjoy your meal’ (literally: Eat them!), *Werk ze!* ‘Have fun working’ (literally: Work them!), described in Cook (1995:226) and Coppen (1998), which cannot be negated either: **Eet ze niet!* ‘Do not enjoy your meal’, **Werk ze niet!* ‘Do not have fun working’.

rather than the verb stem to facilitate that role. Following Contini-Morava's discussion of the asymmetry between positive and negative utterances (1989:3-5, 172-181), one could argue that positive commands, like positive statements, are typically not made against a background of specific negative expectations while negative commands, like negative statements, *are* made against a background of specific positive expectations.¹³ Just as a sentence like *John did not eat the fish* is usually uttered in a context where it might have been possible for John to eat the fish, *Don't eat fish* might be uttered in a context where there would be a possibility that the Hearer would eat fish. Since possibility rather than actuality is usually communicated explicitly, transparently, by modal or at least nonfinite verbal forms, it would make sense that there would be a tendency for prohibitions to be communicated more often with a negator plus IPI rather than a negator plus the more strongly affirmative STM.¹⁴

We conclude by noting that the study of even such apparently humble and particular phenomena as two kinds of imperatives in Modern Dutch and two (out of many) of its pragmatic particles unavoidably touches fundamental theoretical and methodological issues: the distinction between meaning (what is explicitly signaled by a linguistic form) and the messages which such forms are used to communicate, the relative merits of top-down versus bottom-up analyses, the boundary between semantics and pragmatics, and the very nature of commands and negation. We have also attempted to demonstrate that linguistic analysis is

¹³ Consider Contini-Morava's remark (1989:172) that normal discourse is usually more concerned with actually occurring events than with events that fail to occur, whose number is infinite; cf. the oddity of *John did not get up and then he did not have a cup of coffee*, which – though suggesting by mentioning these actions that getting up and having coffee were possible events in the context – still does not tell us what John actually did do. Analogously one could argue that since there are an infinite number of actions which a Hearer might be told not to carry out, commands would be more informative, more communicatively useful if they were employed to tell Hearers what to *do*, what to actually carry out, than what to refrain from doing. This suggests that positive commands will be more frequent than negative ones, which is certainly the case in Table 5, where it is seen that there are $337/58 = 5.8$ times as many positive commands as prohibitions in our data base.

¹⁴ Cf. Hoekstra & Hyams' suggestion (1998:103) that the infinitive specifies the event as [–realized].

served best when it moves beyond isolated anecdotal examples to quantitative data provided by, on the one hand, psycholinguistic experiments with a number of native speakers (such as the questionnaire) and, on the other, detailed studies of actual usage.¹⁵

Acknowledgements

The research reported on here was made possible by Grant 2964 of the Committee on Research of the UCLA Academic Senate. I wish to thank Theo Janssen, Henk Maier, Vincent van Heuven, Rob Goedemans, Jos Pacilly and Johanneke Caspers, among others, for their native speaker intuitions and for numerous discussions of the data. Thanks also to an anonymous referee and to Erica García for critical comments on a previous version. A preliminary, abbreviated discussion of some of this research appeared in Kirsner (2001).

References

- Beets, A. & J.A.N. Knuttel (1912). *Woordenboek der Nederlandsche Taal*. Zesde Deel. Haarst-Izegrim. Martinus Nijhoff, The Hague.
- Blom, A. (1987). KLOPPEN S.V.P. Onderdeel van een procedure. *Voortgang: Jaarboek voor de Neerlandistiek* 8, 177-189.
- Brown, P. & S.C. Levinson. (1987). *Politeness. Some Universals in Language Usage*. (Studies in Interactional Sociolinguistics 4.) Cambridge University Press, Cambridge.
- Buitenhuis, P.A. (1993). *Hoor in vijf gedaanten. Onze Taal* 62.1, 6.

¹⁵ NOTE ADDED IN PROOF. Since this paper was completed and submitted to the editors, two further relevant items have come to my attention. First, in a questionnaire listing various options to select from, I chanced upon the polite instruction *Graag doorstrepen wat niet van toepassing is* ‘Please strike through what does not apply,’ using IPI. Now if STM is ‘gruffer’ than IPI, one would certainly predict that STM would be less compatible with this use of *graag* (literally ‘gladly’) to communicate a request. And indeed, native speakers confirm that *Streep graag door wat niet van toepassing is*, using STM, is quite strange. Second, Egbert Fortuin (2003) “De directieve infinitief en de imperatief in het Nederlands”, *Nederlandse Taalkunde* 9.1.14-43 appeared. This very important article demonstrates that the STM-IPI opposition is even richer – even more subtle and complex – than the initial characterizations of it given by Blom, Duinhoven, and Proeme.

- Contini-Morava, E. (1989). *Discourse Pragmatics and Semantic Categorization: The Case of Negation and Tense-Aspect with Special Reference to Swahili*. Mouton de Gruyter, Berlin and New York.
- Cook, K. (1995). *Dubbel Dutch*. Boekwerk, Groningen.
- Coppen, P.-A. (1998). De succes-imperatief. *Nederlandse Taalkunde* 3, 85-95.
- Duinhoven, A.M. (1984). Ban de bom! Over de vorm en betekenis van de imperatief. *De nieuwe taalgids* 77.2, 148-156.
- Duinhoven, A.M. (1997). *Middel nederlandse syntaxis: synchroon en diachroon. Deel 2: De werkwoordgroep*. Martinus Nijhoff, Groningen.
- Foolen, A. (1995). Dutch modal particles: The relevance of grammaticalized elements. In T.F. Shannon & J.P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1993*, 57-70. (Publications of the American Association for Netherlandic Studies 8.) University Press of America, Lanham, MD.
- Foolen, A. (1996). Pragmatic particles. In J. Verschueren et al. (eds.), *Handbook of Pragmatics 1996*, 1-24. John Benjamins, Amsterdam and Philadelphia.
- Givón, T. (1979). *On Understanding Grammar*. Academic Press, New York.
- Goldberg, A.E. (1992). The inherent semantics of argument structure: The case of the English ditransitive construction. *Cognitive Linguistics* 3, 37-74.
- Haan, S. de (1992). Vul de bon in en win een reis! Over imperatieven in reclameteksten. In E.G. Schermer-Vermeer, W.G. Klooster & A.F. Florijn (eds.), *De kunst van de grammatica. Artikelen aangeboden aan Frida Balk-Smit Duyzentkunst bij haar afscheid als hoogleraar Taalkunde van het hedendaags Nederlands aan de Universiteit van Amsterdam*, 95-109. Universiteit van Amsterdam, Amsterdam.
- Haeseryn, W., K. Romijn, G. Geerts, J. de Rooij & M.C. van den Toorn (1997²). *Algemene Nederlandse Spraakkunst*. 2 volumes. Martinus Nijhoff, Groningen.
- Haiman, J. (1990). Sarcasm as theater. *Cognitive Linguistics* 1, 181-205.
- Han, C. (1998). The Structure and Interpretation of Imperatives: Mood and Force in Universal Grammar. Ph.D. dissertation, University of Pennsylvania.
- Heuven, V.J. van & R.S. Kirsner (1999). Interaction of grammatical form and intonation: Two experiments on Dutch imperatives. In R. Kager & R. van Bezooijen (eds.), *Linguistics in the Netherlands 1999*, 81-96. John Benjamins, Amsterdam and Philadelphia.
- Hoeksema, J. (1992). Bevelende zinnen zonder polaire tegenhanger. In H. Bennis & J.W. de Vries (eds.), *De binnenbouw van het Nederlands. Een bundel artikelen voor Piet Paardekooper*, 125-131. ICG Publications, Dordrecht.
- Hoekstra, T.A. & N. Hyams (1998). Aspects of root infinitives. *Lingua* 106, 81-104.
- Hyams, N. (2002). Child non-finite clauses and the mood-aspect connection: evidence from child Greek. To appear in R. Slabakova & P. Kempchinsky (eds.), *The Syntax, Semantics, and Acquisition of Aspect*. Lawrence Erlbaum Associates, Mahwah, N.J.
- Ijbema, A. (2002). Grammaticalization and Infinitival Complements in Dutch. PhD dissertation, Leiden University. LOT, Utrecht.
- Janssen, T. (1995). Hereosemy or polyfunctionality? The case of Dutch *maar* 'but, only, just'. In T.F. Shannon & J.P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1993*, 71-85. Publications of the American Association for Netherlandic Studies 8. University Press of America, Lanham, MD.

- Kirsner, R.S. (2000). Empirical pragmatics: Downtoning and predictability in a Dutch final particle. In T.F. Shannon & J.P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1997*, 45-62. Publications of the American Association for Netherlandic Studies 12. University Press of America, Lanham, MD.
- Kirsner, R.S. (2001). *De deur wel dichtdoen, hè?* en andere mysterieuze zaken. *Neerlandica Extra Muros* 39.2, 9-21.
- Kirsner, R.S. & J.Y. Deen (1990). *Het mes snijdt aan twee kanten*: On the semantics and pragmatics of the Dutch sentence-final particle *hoor*. In M. Bruijn Lacy (ed.), *The Low Countries: Multidisciplinary Studies*, 1-11. Publications of the American Association for Netherlandic Studies 4. University Press of America, Lanham, Md.
- Kirsner, R.S. & V.J. van Heuven (1996). Boundary tones and the semantics of the Dutch final particles *hè*, *hoor*, *zeg*, and *joh*. In C. Cremers & M. den Dikken (eds.), *Linguistics in the Netherlands 1996*, 133-146. John Benjamins, Amsterdam and Philadelphia.
- Kirsner, R.S. & V.J. van Heuven (1999). How Dutch final particles constrain the construal of utterances: Experiment and etymology. In K. van Hoek, A. Kibrik & L. Noordman (eds.), *Discourse Studies in Cognitive Linguistics*, 165-183. John Benjamins, Amsterdam and Philadelphia.
- Kirsner, R.S., V.J. van Heuven & R. van Bezooijen (1994). Interaction of particle and prosody in the interpretation of factual Dutch sentences. In R. Bok-Bennema & C. Cremers (eds.), *Linguistics in the Netherlands 1994*, 107-118. John Benjamins, Amsterdam and Philadelphia.
- Kirsner, R.S., V.J. van Heuven & J. Caspers (1998). From request to command: An experimental study of grammatical form, intonation, and pragmatic particle in Dutch imperatives. In R. van Bezooijen & R. Kager (eds.), *Linguistics in the Netherlands 1998*, 135-148. John Benjamins, Amsterdam and Philadelphia.
- Lambooy, J.H.Th. (1963). De korte infinitief. In A.W.de Groot (ed.), *Studies op het gebied van het hedendaagse Nederlands*, 107-157. Mouton & Co., The Hague.
- Langacker, R.W. (1991). A usage-based model. In R.W. Langacker, *Concept, Image, and Symbol: The Cognitive Basis of Grammar*, 261-288. Mouton de Gruyter, Berlin and New York.
- Langacker, R.W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (eds.), *Usage Based Models of Language*, 1-63. CSLI Publications, Stanford.
- Levinson, S.C. (1983). *Pragmatics*. Cambridge University Press, Cambridge.
- Osgood, C.E., G.J. Suci & P.J. Tannenbaum (1957). *The Measurement of Meaning*. University of Illinois Press, Urbana, Ill.
- Paardekooper, P.C. (1951). De imperatief als grammatische categorie in het ABN. *De nieuwe taalgids* 44, 97-107.
- Proeme, H. (1984). Over de Nederlandse imperativus. *Forum der Letteren* 25, 241-258.
- Reichenbach, H. (1966). *Elements of Symbolic Logic*. The Free Press, New York.
- Schiffrrin, D. (1987). *Discourse Markers*. Studies in Interactional Sociolinguistics 5. Cambridge University Press, Cambridge.
- Shetter, W.Z. & I. van der Cruysse-van Antwerpen . (2002). *Dutch: An Essential Grammar*. Routledge, London & New York.

- Stoett, F.A. (1923). *Middelnederlandsche Spraakkunst: Syntaxis*. Martinus Nijhoff, The Hague.
- Sweetser, E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press, Cambridge.
- Vismans, R. (1993). Dutch modal particles in directive sentences. In R.S. Kirsner (ed.), *The Low Countries and Beyond*. Publications of the American Association for Netherlandic Studies 5, 111-121. University Press of America, Lanham, MD.
- Vismans, R. (1994). *Modal Particles in Dutch Directives: A Study in Functional Grammar*. Studies in Language and Language Use 11. IFOTT, Amsterdam.
- Weijnen, A.A. (1971). *Schets van de Geschiedenis van de Nederlandse Syntaxis*. Van Gorcum & Co, Assen.
- Weinreich, U. (1957). Travels through semantic space. *Word* 14, 346-66.
- Wouden, T. van der (1998). *Niet huilen maar*: negation, directives and particles. In L.C.J. Barbiers, J.E.C.V. Rooryck & J.M. van de Weijer (eds.), *Small Words in the Big Picture: Squibs for Hans Bennis*, 125-128. Holland Institute of Generative Linguistics, Leiden.
- Zanuttini, R. (1997). *Negation and Structure*. Cambridge University Press, Cambridge.

Sources

- Campert, Remco (1960). *Een ellendige nietsnut en andere verhalen*. De Bezige Bij, Amsterdam.
- Carmiggelt, Simon (1975). *Slenteren*. De Arbeiderspers, Amsterdam.
- Dorrestein, Renate (2000). *Verborgen gebreken*. Pandora Pockets.
- Hellinger (1967). *Vlammen*. Born, Amsterdam.
- Reve, Gerard (no date). *The evenings*. Translated by Paul Vincent. Unpublished manuscript. Foundation for the Promotion of the Translation of Dutch Literary Works, Amsterdam.
- Reve, Gerard Kornelis van het (1972). *De Avonden. Een winterverhaal*. Athenaeum – Polak & Van Gennep, Amsterdam.
- Straten, Peter van (1989). *Lukt het Agnes?* Maarten Muntinga, Amsterdam.

Meaning and use of causeless causative constructions with *laten* in Dutch

Judith Loewenthal (Leiden University)

1. Introduction

The Dutch language user has different ways to express causality. One of these ways is the so called analytic causative construction:

- (1) De leraar laat de kinderen hun huiswerk maken.
The teacher lets the children their homework do
'The teacher makes the children do their homework.'
- (2) De zon doet de temperatuur oplopen.
The sun does the temperature rise
'The sun makes the temperature rise.'

Analytic causative constructions in Dutch are constructions formed by an auxiliary, a form of the verbs *doen* or *laten*, in combination with a bare infinitive. *Doen* or *laten* forms the causal predicate that denotes the underlying cause; the bare infinitive forms the effected predicate that describes the effect or consequence. The analytic causative construction (from now on CC) has at least two participants: a *causer*, which is the subject of the construction, and a *causee*, which is the participant that 'carries out' the effected predicate.

The difference in form between the *doen*- en *laten*-construction corresponds to a crucial difference in meaning (Kemmer & Verhagen 1994; Verhagen & Kemmer 1992, 1997; Verhagen 2000). With *doen* the construction expresses direct causality: the effect of the causal event is inevitable, the result is the inevitable consequence of the underlying causing event. *Laten* marks indirect causality: the causing event is not a sufficient condition for the realization of the effected predicate, but there is another force active that is more directly involved in the realization of the effect. Both direct (with *doen*) and indirect causality (with *laten*) can be further divided in a number of causality subtypes. (Verhagen &

Kemmer 1992, 1997; Kemmer & Verhagen 1994; Talmy 1988; Croft 1991). These subtypes are related to the participants of the constructions and their properties, especially their animacy or inanimacy. Example (2) with inanimate causer and inanimate causee expresses physical causation, a subtype of direct causation. Example (1) with *laten* expresses indirect, inductive causation: both causer and causee are animate participants.

Indirect, inductive CCs (constructions with *laten*, and animate causer and causee) can be further categorized at a third level. Their meanings can be divided in permissive causation, meaning ‘to permit’ or ‘not hinder’ (contributing to the realization of an event by not interfering, (example (3)), and coercive causation, meaning ‘to cause’, ‘to force’ or ‘to coerce’ (example (1)).

- (3) De conducteur liet de hooligans de trein afbreken
The conductor let the hooligans the train demolish
‘The conductor let the hooligans demolish the train.’

The difference between both meanings can be described in terms of the conceptualization of the roles of the causer and the causee in the overall causal event. If the realization of the effected predicate is in accordance with the ‘wish’ of the causee, and the causer is neutral in this respect, then the construction gets a permissive meaning. In a prototypical coercive construction on the other hand, the causee carries out the effected predicate against his will, and is forced to act by the causer. Below, the differences between the constructions and the roles of the participants will be defined more precisely.

In some uses of the *laten*-CC, the causee remains unexpressed, as in (4) (the *affectee* is the object of the effected predicate):

- (4) De huiseigenaar [causer] laat zijn huis [affectee] overschilderen.
The home-owner lets his house repaint
‘The home-owner has his house repainted.’

The possibility to leave the causee implicit has first of all to do with the transitivity of the verb in the effected predicate: CCs with a transitive effected predicate and an explicit affectee can leave the causee unexpressed.

This property of the CC has often been described in the literature (Verhagen & Kemmer 1994; Degand 2001), but never fully in combination with the semantics of this causelessness. Which CCs actually do leave their causees implicit? It seems difficult for instance to conceptualize the expression of permission without the presence of a causee. In the light of Talmy's Force Dynamics (1988) it seems plausible to state that the force of the causee is a crucial force in the realization of the effected predicate in a permissive construction. The causer does not move in the direction of the causee, to make the causee realize the effected-predicate, but makes a way for the causee to carry out the effect by moving away. The causer leaves the scene so to speak, and by doing so he gives the causee the opportunity to continue his route. Because of this necessary force of the causee in the conceptualization of permission, one would expect the causee always to be expressed explicitly in a CC with permissive meaning. This in contrast to constructions that express coercive causation, where the force of the causer moves in the direction of the effected predicate, and where the opposing force of the causee is being overruled by the force of the causer. In these cases it should be possible to leave the causee implicit, as in example (4) above and also in (5):

- (5) Hij [causer] laat een omheining [affectee] om zijn tuin bouwen.
 He lets a fence around his garden build
 'He has a fence built around his garden.'

In other words, there has to be some relation between the non-expression of the causee of the construction on the one hand, and the interpretation of that construction on the other hand. The prediction that removal of the causee of the construction relates to a causal, non-permissive interpretation seems plausible.

The aim of this paper is to investigate the most important uses of the causative construction with *laten*, and to investigate which *laten*-CCs do leave their causees implicit and which do not. The expectation is that constructions with no explicit causee are less suitable for the expression of permission. On the other hand, causeless constructions are well suited for the expression of coercive causality. Testing these expectations and the investigation of the uses of the *laten*-construction will be done by means of corpus analysis, and by analysis of actual occurrences of the *laten* causative constructions. My analysis will show that the corpus data

fit the expectations in broad outlines. However, there are some identifiable groups of constructions whose behavior deviates systematically from the expectations. These groups of divergent constructions seem to have their own specific features, grammatically and semantically.

2.1 Interaction between participant semantics and overall construction meaning

In order to understand the consequences of the absence of the causee and the sole presence of the causer for the semantics of the different causative constructions, we have to understand the different roles the causer and the causee play in the different causal events expressed by the CCs. The causer is the subject of the construction and the underlying source of the causal event. The causee is the agent of the effected predicate, the actor that “carries out” the effect. The causer “does something”, expressed by the causal predicate, as a result of which the causee carries out the effected predicate. When the language user uses the CC, he presents (some action of) the causer as responsible for the realization of the effect by the causee. Degand formulates the roles of causer and causee in terms of control:

“Thus it seems that the Agent of the underlying causing event, which we will label causer [...], is given a special status in the overall causative situation: that of (a certain degree of) control of the situation, and especially over the agent of the underlying caused event, labeled the causee [...]. In other words, a speaker will use this type of construction only if he/she believes that the Causee is under control of the Causer or plays a minor role in the causal event.” (Degand 2001:176)

Kemmer & Verhagen (1994) explain the roles of the participants and the meaning of the construction in terms of transfer of energy: the causer is the source of energy, the causee absorbs this energy and carries out the effect. When there is an affectee present, the object of the transitive effected predicate, the stream of energy will be extended:

“The causer [...] is the entity viewed as causing the entire event [...]. The causee is the entity carrying out the activity designated by the effected predicate [...]. The Affectee, where present, is the entity that is the

endpoint of energy (literal or metaphorical) expended in the entire causative event [...]” (Kemmer & Verhagen 1994:119).

“We use the term ‘affectee’ because the participant in question is in the most prototypical cases affected by the causal event” (Kemmer & Verhagen 1994:149).

We have to further elaborate the description of the participants by looking at the different causality types that the CCs express, and the roles the participants play in the realization of the effected predicates in the different CCs. It is useful in this respect to look at the difference between direct causality (*doen*) and indirect causality (*laten*) in terms of the roles of the participants. In *direct* causation, the causer’s role is maximal, while the role of the causee is minimal. The causer’s energy is enough to produce the effect, although he is not the one who actually carries out the effected predicate. Because the action of the causer is seen as a sufficient condition to realize the effect, the causee’s role can be seen as minimal and not relevant in the realization of the effect, although the causee is in fact the one that “carries out” the activity. There is no intention needed from the side of the causee to carry out the effect, the effect happens beyond his consciousness or control. In *indirect* causation, the causer’s role is not a sufficient condition for the realization of the effect although he still is the source of the causal event. There is another force present that is more directly involved in the production of the effect. This can be a third force, as in (6), where it is gravity that is more directly involved in the water’s flowing out of the bathtub. The force can also be inside the causee, as in (7), where it is the causee itself that makes the last decision in carrying out the activity:

- (6) Hij laat het badwater weglopen.
He lets the bath water out flow
‘He lets the water flow out of the bathtub.’
- (7) Hij laat de kinderen het plein op rennen.
He lets the children the square up run
‘He lets the children run onto the square.’

All causees of indirect constructions have some degree of autonomy (Kemmer & Verhagen 1994; Stukker et al 1999; Degand 2001), that

enables them to move to some degree according to their own 'wish'. For animate causees, this autonomy is closely related to volition (Stukker et al. 1999), the will to carry out the effected predicate or not. In the end, the causees need to have at least some intention to carry out the effect. For inanimate causees the autonomy is related to the already mentioned third force, a force of nature, like for instance gravity.

The degree of autonomy of causees is variable, and differs according to the different subtypes of indirect causation. Since the causers in inductive constructions are animate, they too have some mental stance with respect to the effected predicate: the effected predicate can be their volitional goal, or it may be just the (unintended) consequence of their actions. Let us look at the roles of the participants in the following constructions:

- (8) De leraar laat de kinderen de hele middag buiten spelen.
The teacher lets the children the entire afternoon outside play
'The teacher lets the children play outside for the entire afternoon.'
- (9) De moeder laat de kinderen vroeg naar bed gaan.
The mother lets the children early to bed go
'The mother makes the children go to bed early.'
- (10) Hij zal het ons laten weten.
He will it us let know
'He will let us know.'

The example in (8) is of the subtype permissive causality: the teacher allows the children to play outside for the entire afternoon. The causees are autonomous to a high degree and act according to their own wishes. The goal of the causer does not seem to be the children's playing outside per se. The playing is just a consequence of his actions. (9) expresses coercive causality: the mother forces her children to go to bed. The causees are not very autonomous, carrying out the effected predicate probably does not correspond to their own wishes. Goal of the causer on the other hand is the realization of the effected predicate by the children. The meaning of example (10) is somewhere between permission and coercion. It is not clear whether we have to do with a causer giving permission or not. Is it the causer's wish, or is it the wish of the causee that 'we' know? It does not seem to be either a matter of coercion and force,

or a matter of permission in which the causer has to move a barrier. The causer takes care that the causee gets certain knowledge, he is going to report to 'us' on something. This example is an 'in-between-form' of inductive causation, where "it does not make much sense to try to decide between a permissive and a causative reading: the use of *laten*, categorizing an event as involving indirect causation, in itself simply leaves this undecided" (Verhagen & Kemmer 1997: 97). There are more of these 'in-between-forms', which cannot be categorized as either permissive or coercive.

Permissive meaning, as compared to coercive meaning, is the older meaning of the *laten*-construction. So-called in-between-cases did exist next to the permissive constructions in a period that the *laten*-construction did not express coercion yet. These in-between cases were mostly CCs with perception verbs and cognition verbs as effected predicates. Examples of these verbs are: *zien* 'to see', *horen* 'to hear', *merken* 'to notice', *lezen* 'to read' en *weten* 'to know' (Dik 1980). In modern Dutch, constructions with these effected predicates have a meaning in between permission and coercion too:

- (11) Kort voor het jaareinde hebben de politieke leiders van het land elkaar nog even laten weten hoe zij over elkaar denken. (EC: 2484)
 Shortly before the year end have the political leaders of the country each other just let know how they about each other think
 'Shortly before the end of the year the political leaders have let each other know how they feel about each other.'
- (12) [...] om hen in een werksituatie te laten merken dat het samen functioneren iets anders kan zijn dan alleen maar iets samen produceren. (EC: 2989)
 [...] to them in a work situation let notice that the together functioning something different can be than just something together produce
 '[...] to let them notice in a work situation that working together can be different from just producing something together.'

- (13) “Als je dit mooi vindt, moet je ook Dean Martin en Frank Sinatra mooi gaan vinden” zei iemand die ik het liet horen. (EC: 2526)
 ‘ “If you this like, have you also Dean Martin and Frank Sinatra go like” said someone that I it let hear
 ‘ “When you like this, you also have to like Dean Martin and Frank Sinatra,” said someone I played the music for.’

Effected predicates with perception verbs and cognition verbs seem to give the participants in the CCs specific roles, characteristic for constructions with just these effected predicates. The causees seem to be more or less passive, they do not have to ‘act’, but are just experiencers. The activity comes from the causer: with the effected predicate as his goal, he has to act to make it possible for the causee to realize the effect. Example (12) for instance, has an acting causer that wants the effected predicate to happen; the causee has a neutral wish and can stay passive. Example (11) is almost similar in this respect. The causers inform the causees on something; that is their goal, they perform the action of telling something. The causees are passive: they do not act but are just auditors, maybe even against their will.

On the one hand, one could speak about removing of a barrier by the causer, as in permission, but on the other hand, the realization of the effected predicate is the goal of the causer, unlike in permission, where it is more the consequence of the causer’s actions. The causee’s wish is not really to carry out the effected predicate, as it is in permissive causation. The autonomy of the causees is not very high either, the effected predicate can be realized without the causees’ volitional actions. Looking at these characteristics, the constructions seem to have properties of permissive constructions and of coercive constructions, but also properties of their own.

So far, we have looked mainly at the interaction of the different semantic factors within the causative constructions, and at the interaction between the semantic factors and the overall construction meaning. In the next section we will look more closely at how the different causality subtypes expressed by the constructions can be related to each other, in terms of degree of indirectness of causation (Kemmer & Verhagen 1994), and at possible implications thereof for the use of the different constructions, especially for the expression of the causee.

2.2 Continuum of indirectness

Verhagen and Kemmer say the following about the semantic relation between the indirect permissive and coercive constructions and the other inducive constructions (Verhagen & Kemmer 1997:69):

“In fact, what we find with *laten* is a continuum of indirectness. At one end (enablement/permission), there are cases with relatively great autonomy of the causee and some inherent tendency for it to carry out the effected predicate; [...]. At the other end (coercive-causative) are cases in which those factors are much more reduced and the causee would not normally carry out the effected predicate in the absence of an external force [...]. There are also in-between-cases.”

Laten-CCs can be placed on a continuum of indirectness on the basis of their overall meaning, which is based on and related to some features of parts of the construction, like participant roles. According to the quotation above, the notions indirectness, permissive causality and coercive causality are related as follows: the degree of indirectness of the construction is related directly to the degree of autonomy of the causee. At one end of the continuum, the maximal indirectness end, permissive CCs can be found: constructions with the most autonomous causees. At the other end of the continuum, the minimal indirectness end, we can find the coercive CCs: constructions with the least autonomous causees. In between these permissive and coercive CCs, in the middle of the continuum, we find all kinds of ‘in-between-cases’. In other words: the more autonomous the causee, the more permissive the construction, the more indirect the (causation expressed by the) construction.

According to this analysis, there is a clear relation between characteristics of the causee of a construction, and the place of that construction on the indirectness continuum. The question is: how can we fit constructions with no causee into this analysis? Where – and how – do we place constructions with no causee on the continuum; and how autonomous is an implicit causee? Let us look at some examples of causee-less causative constructions:

- (14) Ik laat de was doen.
 I let the laundry do
 ‘I have the laundry done.’

- (15) Ik laat de fiets maken.
 I let the bike repair
 'I have the bike repaired.'

Kemmer and Verhagen give the following description of causeless constructions (Kemmer & Verhagen 1994:139):

“[...] *Zij laten een huis bouwen* ‘They are having a house built’, [...]. In our view, causeless causatives constitute the extreme case of peripherality of the causee participant: here, rather than being elaborated in the clause with specific semantic content, the causee remains entirely schematic and is simply not expressed.”

An implicit causee can be seen as an extremely peripheral causee, according to this quotation. Highly relevant now is the question how we can relate peripherality of the causee to autonomy of the causee (and thus to the degree of indirectness):

“Our claim is that case-marked participants in general differ in degree of conceptual integration in the clausal event, with accusative most integrated, dative less integrated, and instrumental and or agentive least integrated, where high degree of integration correlates with high degree of affectedness and topicality and low degree of autonomy of the causee” (Kemmer & Verhagen 1994:138).

If anything, Kemmer and Verhagen state that high degree of integration and high degree of topicality of the causee correlate with low degree of autonomy of the causee. High degree of peripherality of the causee, as in causeless constructions, means absence from the event and no integration in the event and thus a non-affected autonomous causee. Highly autonomous causees, as we saw before, are causees of permissive constructions, so causees of permissive constructions should be good candidates for staying implicit. However, as we have seen, the initial expectation in this paper was just the opposite, namely that causees in permissive constructions, i.e. highly autonomous causees, are most likely expressed explicitly, and constructions with no causee are most likely to have a coercive meaning.

In the light of these conflicting expectations, and in order to investigate the most important uses of the *laten* CCs, we have to look at

actual occurrences of the causative constructions, and at the semantics of (causeeless) causative constructions as they can be found in actual language use.

3. Corpus analysis

I used corpus analysis to investigate the uses of the *laten* CCs, because corpus data will give the best insights in the form-meaning relations in real occurrences of this *laten*-causative construction. The corpus I used is the Eindhoven Corpus (Uit den Boogaart 1975), the same corpus that Kemmer & Verhagen (1992, 1994, 1997) used for their analysis of Dutch causative constructions. This corpus should give a representative picture of modern, mostly written Dutch. I collected only *laten* CCs from the corpus that are relevant for this survey. Because I want to see relations between causelessness and permission and coercion, only the constructions with animate causers and causees (inducive constructions) are relevant, since only these constructions allow permissive as well as coercive readings. From the inducive constructions we have to select the constructions with a transitive effected predicate and an affectee, because only these constructions have the possibility to leave the causee unexpressed. I looked at constructions with explicit causee and without explicit causee, so the semantics of both constructions can be compared.

The CCs that were selected on base of the above mentioned criteria were divided in three groups, according to their semantics: permission, coercion and 'in-between'; the judgements were checked by another native speaker, and disagreements were resolved by discussion. The 'in-between' group consists of the constructions with perception and cognition verbs in their effected predicates which cannot be called either permissive or coercive. Finally there is a fourth category, consisting of constructions that do not fit in one of the three other categories.

3.1 Results

Table 1 shows the results for the causeless constructions, compared with the results for constructions with three explicit participants. The first column gives the results for the CCs with explicit causee, the second column shows the results for the causeless constructions, and the third column gives the total of constructions in the different categories.

	<i>3 participants</i>		<i>Causeless</i>		<i>Total</i>	
Permissive	12	12%	18	11%	30	11%
In-between	31	31%	41	25%	72	27%
Coercive	44	44%	84	52%	128	49%
Rest	13	13%	20	12%	33	13%
Total	100	100%	163	100%	263	100%

Table 1. *Distribution of inductive CCs (with laten) with three participants and without causee*

At first sight, there seems to be no clear difference between CCs with explicit causee and causeless CCs with respect to the expression of permissive causality. The causeless constructions allow the expression of permission despite the expectations. Permission is not frequently expressed by either one of the constructions at all. The results in the third row show that coercive causality is expressed frequently by both types of constructions, with a preference for the causeless one. The results of the ‘in-between’-category given in the second row show that both constructions are well suited to express causality when they have a perception or cognition verb as their effected predicate.

The fourth row in the table mostly consists of *negative* CCs. The reason these constructions are categorized separately, is that negation seems to give the constructions ambiguous meanings. Compare the following examples (16) without negation and (17) with negation:

(16) Ik laat hem gaan.
‘I let him go’

(17) Ik laat hem niet gaan.
a. ‘I do not let him go’
b. ‘I make him stay’

(16) is just permissive: ‘he’ wants to go, and ‘I’ take away the barrier so ‘he’ actually can go. Example (17) can get a (negative) permissive interpretation: the ‘he’ wants to go and ‘asks’ for permission to do so, but the ‘I’ refuses to remove the barrier (a). The other possible interpretation is that the causer creates a barrier which *forces* the causee to stay in his

place, a more causal interpretation (b). This ambiguity makes it difficult to put these constructions in one category, so they are put in a separate category. 13% of the instances got a place in this category. In the next sections we will look at the results from Table 1 in more detail.

3.2 Coercive constructions

The expectation about the semantics of causeless constructions was motivated by the conceptualization of the roles the participants have in the events expressed by the constructions: coercive constructions have little autonomous causees, who act according to the wishes of the causer and whose own force is being overruled. This led to the hypothesis that coercive causality can be expressed especially easily by causeless constructions.

This expectation about the suitability of causeless constructions for the expression of coercion seems to be right: 84 of the 128 instances have an implicit causee. However, coercion is also frequently expressed by the CC with three explicit participants. Because the two constructions both express coercive meaning frequently, it is interesting to see if there is a difference in meaning between the causeless coercive constructions and the coercive constructions with three explicit participants, since a difference in form may be expected to correlate with some difference in meaning.

Verhagen and Kemmer give the following characterization of the causeless CC (1997:63):

(‘Hij heeft een rolstoel laten bouwen’)

“This sentence means that he caused someone, who is not identified any further, to build a wheelchair: thus the interpretation of the causee, being left out, is highly schematic: because the causee receives no independent linguistic expression, its interpretation is exhausted by the information provided by the effected predicate, which evokes the role of a ‘builder’ (of a wheelchair).”

The fact that the causee is not further specified implies that the identity of the causee is not relevant or unexpected. This in combination with the notion coercive causation leads to an interesting expectation as to which constructions have no explicit causee; they should be constructions where the causee has some tendency to carry out the effect, for instance,

constructions where the causee has the effected predicate as his job or profession. When we take a close look at the roles of the participants in the causeless CCs, the difference between the causeless (coercive) CCs and the CCs with explicit causee becomes visible.

The causer wants the effect expressed by the effected predicate to happen, his goal is that the wheelchair will be built. Obviously, the identity of the builder does not really matter. This causee, although not expressed, can be interpreted as a participant who is not unwilling to build the wheelchair, it might even be his job. He does not have to be forced to come into action, he just has to be told to do so (get an order from a client). We do not expect that there will be pressure involved. It is important to notice, however, that the causee is not going to act without a request of the causer.

Looking at the actual data from the Eindhoven Corpus, the coercive causeless constructions all seem to be precisely this kind of events, where the causee does the effected predicate more or less as its job or profession and where the causer's sole goal is the effected predicate to happen, regardless of the identity of the causee. Below are some examples:

- (18) Op een straathoek kan men nog steeds zijn schoen laten verzolen.
(EC: 2727)
At the corner of the street can one still his shoes let resole
'One can still have his shoes resoled at the corner of the street.'
- (19) Ik heb het buitenverblijf laten afschermen. (EC: 2886)
I have the countryhouse let screen
'I have had the countryhouse screened.'
- (20) Ik zal in ieder geval laten zingen psalm 119, vers 18. (EC: 2772)
I will in any case let sing psalm 119, verse 18
'I will in any case have psalm 119, verse 18 sung.'

Examples (18) and (19) describe events where the causer can indeed be interpreted as a participant who makes a request or can make a request for the effected predicate to be acted out, and where the causee does the action expressed by the effected predicate as a job. In example (20), the causer also makes a request, but in this case, the effected action, the singing, does not have to be a real job or profession. It is the institutional

role of the church choir, or the church community (or whatever the appropriate way to interpret the causees) to sing in the church. In this way their role is still closely related to a job or a profession.

The causer's goal in these CCs is the realization of the effected predicate only, the identity of the causee does not seem to matter, as it does in example (21), where it is highly relevant that it is the children that brush their teeth and not somebody else. The causer also expects that the effect will be carried out without any force. Therefore it is not surprising that the causee does not have to be expressed in the examples above.

It must be clear by now that the causeeless coercive constructions do not find themselves at the extreme end of the continuum of indirectness. The meaning of these constructions is not extremely coercive: the (implicit) causees are not forced by the causer to act out the effected predicate. The identity of the causees does not seem to matter, as long as the effected predicate will be realized. This in contrast to coercive constructions with explicit, non case-marked (i.e. zero-marked) causees (see Kemmer & Verhagen, 1994), like the following:

- (21) De moeder laat de kinderen hun tanden poetsen.
The mother lets the children their teeth brush
'The mother lets the children brush their teeth.'
- (22) U kunt ook uw man eens de boodschappen laten doen voor het weekend. (EC: 2707)
You can also your husband sometime the shopping let do for the weekend
'You may also have your husband do the shopping for the weekend.'

In these examples the identity of the causee is important. Example (21) does not even allow the causee to stay implicit. Example (22) gets a totally different meaning without causee, namely a meaning that is characteristic for the coercive causeeless CC, as was described above. This example also demonstrates the ambiguity that constructions with a zero-marked (non-case marked) causee sometimes can show: both a permissive and a coercive interpretation seem to fit in this example (see also example (1)).

So, coercive constructions do allow implicit causees: quite a large amount of the coercive constructions from the corpus have implicit causees. Since coercive constructions do not have highly autonomous causees, peripherality (implicitness) of the causee cannot correlate with autonomy of the causee. On the other hand, coercive meaning of the CC is not a sufficient condition for leaving the causee implicit: not all coercive meanings can be expressed with a causeeless construction (for instance, causal events in which causees are forced to realize the effect). Causeeless coercive CCs form a subgroup of coercive CCs and cannot be found at the extreme end of the indirectness continuum, where the explicit, zero-marked causees are found. They express a more moderate coercive meaning, and should be placed more towards the middle of the continuum.

3.3 Permissive constructions

Causees in permissive constructions are more autonomous than coerced causees are. Carrying out the effected predicate is the wish of the causee, and in order to do so, the causer has to leave the stage, thereby removing a barrier (Talmy 1988). The causee has to complete the effect on his own force so to speak. Because of this necessary force of the causee in the realization of the effect, the prediction was made that permissive constructions will not leave their causees implicit, and that causeeless constructions will probably not get a permissive interpretation.

However, from Table 1 it can be seen that there are causeeless constructions with permissive meaning: 18 of the 30 constructions. However, these constructions share a striking similarity that sets them apart from the other constructions found in the corpus: the permissive constructions that do not have explicit causees are all *reflexive* CCs: causer and affectee refer to the same entity. In the next section, I will show that the original hypothesis still holds: permissive constructions normally do not leave their causees implicit, but there are specific contexts that allow exceptions.

3.3.1 Middle voice

Some examples of the permissive causeeless constructions that are found in the data are the following:

- (23) Het lagerhuis eiste in 1621 dat hij in staat van beschuldiging zou worden gesteld, omdat hij [causer], naar het heette, zich [affectee] had laten omkopen. (EC: 2960)
 The Lower House demanded in 1621 that he in state of charge would be put, because he, as it was called, himself had let be bribed
 ‘In 1621 the Lower House demanded that he would be indicted, because he had excepted bribes.’
- (24) Men bestreed Hannah Arendts [...] ideeën over de houding van het joodse volk – waarom hebben de joden zich als schapen naar de slachtbank laten voeren? (EC: 2474)
 People fought Hannah Arendt’s [...] opinions about the attitude of the Jewish people – why have the Jews themselves like sheep to the slaughter let carry?
 ‘People disagreed with Hannah Arendt’s [...] view on the behavior of the Jewish people – why did the Jews let themselves be slaughtered like sheep?’
- (25) Want de landbewerkers laten zich niet langer zo bedriegen. (EC: 2496)
 Because the farmers let themselves no longer so deceive
 ‘Because the farmers won’t let themselves be deceived like that any longer.’

These examples, like the other causeless permissive constructions from the corpus, are *reflexive* constructions: causer and affectee refer to the same entity. Since “reflexive” constructions seem to have a special status in the data, I will look at their semantics and form in closer detail. Reflexive constructions with explicit causees will also be included in this analysis. Several kinds of CCs can be called reflexive, because different participant combinations can be coreferential. Consider example (26); here causer and *causee* are referring to the same entity:

- (26) Ik laat me gaan.
 ‘I let myself go’

In my data, the relevant coreference is the type occurring in examples such as (27); the coreferential participants are causer and *affectee*:

- (27) Ze verweten hem niet dat hij zich door ontactisch optreden uit de regering had laten verdrijven. (EC: 2414)
 They did blame him not that he himself by tactless behavior out of the government had let driven
 ‘They did not blame him for letting himself be excluded from the government because of his undiplomatic behavior.’

In the literature, not much can be found about the semantics of these constructions in Dutch (I will call these CCs “reflexive causative constructions” for the moment). Dik (1980) mentions the frequency of occurrence of the reflexive causatives, not only with actual coreference between causer and affectee, but also with other kinds of coreferential relations, like possessive relations. Concerning the semantics of this reflexive CC, the *Algemene Nederlandse Spraakkunst* (ANS) says the following:

“Bij dit laatste type [causatiefconstructie in wederkerende verbinding] kan opgemerkt worden dat de infinitief bij *laten* soms passieve betekenis heeft. Bevat de zin een door-bepaling [causee verbonden met door] dan is de betekenis “toelaten” nog duidelijk aanwezig [...]. Zonder door-bepaling kan de betekenis van *laten* + infinitief naderen tot: “...kunnen worden”.” (ANS 1983:187)

[Translation: For this last type [reflexive CC] it can be noticed that sometimes *laten* + infinitive has passive meaning. When the clause contains a causee with *door*, then the meaning “to allow” is still clearly present [...]. Without the presence of a causee with *door*, the meaning of *laten* + infinitive can approach the meaning “can be.”]

A paraphrase with ‘can be’, however, seems to be more suitable for inanimate CCs in which causer and affectee refer to inanimate entities, as in (28) and (29):

- (28) Rasproblemen laten zich moeilijk van verre beoordelen en oplossen. (EC: 35885)
 Racial problems let themselves hardly from a distance judge and solve
 ‘Racial problems are hard to be judge and hard to solve from a distance.’

- (29) Gezouten vlees laat zich langer bewaren. (ANS: 187)
 Salted meat lets itself longer conserve
 ‘Salted meat can be kept longer.’

Kemmer (Kemmer 1993), mentions these so-called passive-middle constructions, the semantics of which show similarity with the *facilitative* (Kemmer 1993:147), but which differ from constructions with animate causers in a number of ways.

In *inanimate* reflexive constructions, the emphasis is on some property or quality of the causer (and affectee), whereas reflexive CCs with animate causers refer to causal events in which the causer participates. *Inanimate* reflexive CCs do not have an explicit causee, but an implied causee which can be paraphrased as ‘people, in general’ (Fagan, 1992: 211). Many animate reflexive constructions on the other hand, do have an explicit causee, and when they don’t, the implied causees have specific interpretations.

Although the *Algemene Nederlandse Spraakkunst* suggests the same paraphrase for both the animate and the inanimate reflexive construction, the animate one does not seem to fit the paraphrase. Example (30a), for instance, cannot be just paraphrased with ‘can be’, as in (30b):

- (30) a. Ik betreur het, dat behalve de KNVB ook de FBO zich heeft laten intimideren. (EC: 2680)
 I regret it, that apart from the KNVB also the FBO himself has let intimidate
 ‘I regret that, besides the KNVB, the FBO also let itself be intimidated.’
- b. Ik betreur het, dat behalve de KNVB ook de FBO geïntimideerd heeft kunnen worden
 ‘I regret that, besides the KNVB, the FBO could also be intimidated.’

It should be clear that there is more to the meaning of (30a) than is present in a paraphrase with ‘can be’.

The other meaning giving by the ANS, ‘to allow’, is the same as the meaning of a normal prototypical, non-reflexive permissive CC. Although ‘to allow’, and the related meanings ‘to permit’ and ‘not-

interfere' can be used as a paraphrase of a normal permissive CC (see for instance (31a) and (31b)), they cannot cover the entire meaning of the (permissive) reflexive constructions. See for instance example (32a) and its paraphrase with 'to permit' in (32b).

- (31) a. Hij laat de kinderen de klas uitrennen.
 He lets the children the classroom out run
 b. Hij staat toe dat de kinderen de klas uitrennen.
 'He permits the children to run out of the classroom.'
- (32) a. Hij liet zich misbruiken.
 He let himself abuse
 b. Hij stond toe dat hij misbruikt werd.
 'He permitted himself to be abused.'

Permission is an important part of the semantics of the reflexive CCs, but there is more to their meaning, which becomes clear when we go back to the general characterization of causative constructions and their participants as was given by Kemmer and Verhagen:

"The causer [...] is the entity viewed as causing the entire event [...]. The causee is the entity carrying out the activity designated by the effected predicate [...]. The Affectee, where present, is the entity that is the endpoint of energy (literal or metaphorical) expended in the entire causative event [...]" (Kemmer & Verhagen 1994:119).

"We use the term 'affectee' because the participant in question is in the most prototypical cases affected by the causal event" (Kemmer & Verhagen 1994:149).

For reflexive causal events this means that the entity causing the entire event is the same as the endpoint of energy. In other words, the causer is affected by the causal event from which he himself is the source. And so the causer, being the same entity as the affectee, is this time self affected by the causal event over which he has (some) control. One could say that he is 'responsible' for his own affectedness. When this is combined with the notion permissive causation the picture becomes even clearer: the causer *allows* someone, or some force to affect him, he *allows* himself to be affected.

A causee in a prototypical permissive construction is a participant that wants the effected predicate to be realized. We might consequently hypothesize that effected predicates of prototypical permissive constructions are in the first place in the interest of the causees. Let us look at the effected predicates of the permissive reflexive constructions. From the 36 effected predicates (including those with explicit causee and with negation), 33 involve a clearly negative meaning, two are neutral in this respect and only one can be called positive:

Negative meaning

- met zich laten spotten ('to let o.s. be mocked')
- zich (naar de slachtbank) laten voeren ('to let o.s. be led to the slaughter')
- zich de kaas van het brood laten eten (2x) ([fig.:] 'to let o.s. be bullied')
- zich iets laten doen ([lit.:] 'to let o.s. be done something')
- zich iets laten gezeggen ([fig.:] 'to let o.s. be said what to do')
- zich iets laten verbieden ('to let o.s. be forbidden something')
- zich iets laten zeggen ([lit.:] 'to let o.s. be said something to')
- zich in de kaart laten kijken (3x) ([fig.:] 'to give one's hand away')
- zich in een hoek laten drijven ([lit.:] 'to let o.s. be forced into a corner')
- zich in een hoek laten drukken (idem.)
- zich in een positie laten manoeuvreren ('to let o.s. be manoeuvred into some position')
- zich laten bedriegen (2x) ('to let o.s. be fooled')
- zich laten bezwadden ('to let o.s. be defiled')
- zich laten intimideren (2x) ('to let o.s. be intimidated')
- zich laten kisten ([lit.:] 'to let oneself be boxed')
- zich laten leiden (door de duivel) ('to let o.s. be led by the devil')
- zich laten misbruiken ('to let o.s. be abused')
- zich laten omkopen ('to let o.s. be bribed')
- zich laten ompraten ([fig.:] 'to allow o.s. to be brought around')
- zich laten schilderen
(the complete sentence is '*Laat jij je zo schilderen?*' had zij ontsteld aan
Carla gevraagd ('to let o.s. be painted'))
- zich laten uitvreten ('to let o.s. be sponged on')
- zich laten verdrijven ('to let o.s. be excluded')
- zich onder druk laten zetten ('to let o.s. be put under pressure')
- zich op sleptouw laten nemen ('to let o.s. be led by')
- zich van de bal laten dringen ([lit.:] 'to let o.s. be pushed away from the [soccer-] ball')
- zich laten fêteren ([lit.:] 'to let o.s. be lionized')
- zich laten verleiden ('to let o.s. be seduced')
- zich laten bepraten ('to let o.s. be talked into something')

neutral meaning

zich laten behandelen ('to let o.s. be treated')

zich laten bezoeken ('to let o.s. be visited')

positive meaning

zich laten vertroetelen ('to let o.s. be pampered')

From this table it can be concluded that reflexive permissive constructions have a striking preference for negative effected predicates. This means, that there has to be at least an association with a negative meaning. When this pejorative sense is combined with the notion of permission or allowance and the fact that in reflexive CCs the causer and causee are the same entity, we get the following. First, the strong correlation between permission and a negative effect is clearly motivated: one normally tries to prevent negative results. Second, the causer in a reflexive permissive construction allows himself to be affected (by someone or some force) in a negative way, and allowing yourself to be affected negatively, is negative in itself: Instead of protecting yourself or defending yourself, you allow someone to affect you in a negative way. By using this construction, the speaker can show his opinion that the causer is (in some way) responsible for his own state, and that he should have acted differently from what he did: he should have had protected himself.

An illustration of this meaning is the following view on the holocaust:

- (33) Men bestreed Hannah Arendts ideeën over de houding van het joodse volk – waarom hebben de joden zich als schapen naar de slachtbank laten voeren? [EC: 2474)
 'People disagreed with Hannah Arendt's [...] view on the attitude of the Jewish people – why did the Jews let themselves be slaughtered like sheep?'

From the use of the reflexive permissive CC in this context, we can infer that Hannah Arendt blames the Jews in a way for not defending themselves. Because they did not defend themselves against the Germans, they created an opportunity for the Germans, and contributed in a way to

the realization of the effect, while they should have done something, they should have stood up for themselves. If the speaker does not want to give some responsibility to the subject for his state, his negative affectedness, he will use a construction that is neutral in this respect, for instance a passive construction, instead of a causative construction.

3.3.2 Middle events

Reflexive constructions have yet another remarkable property. Affectees of ‘normal’ causative constructions can easily be replaced by another entity and still be part of a well-formed sentence, but affectees of permissive reflexive CCs cannot be so replaced. In *permissive* reflexive constructions, the coreference between causer and affectee has to stay intact, in order to have a well-formed sentence:

- (34) a. Ik betreur het, dat behalve de KNVB ook de FBO zich heeft laten intimideren [...]. (EC: 2680)
 ‘I regret it that, apart from the KNVB, the FBO has also let himself be intimidated.’
 b. [?] Ik betreur het, dat behalve de KNVB ook de FBO de spelers heeft laten intimideren
 ‘I regret that, apart from the KNVB, the FBO also let the players be intimidated.’
[?] ‘I regret that, apart from the KNVB, the FBO also let the players be intimidated.’
- (35) a. De grote Rik van Steenbergen, [...], heeft zich als een kleine jongen laten misbruiken als opiumkoerier. (EC: 2418)
 ‘The great Rik van Steenbergen [...] has let himself be exploited as drugs courier.’
 b. [?] De grote Rik van Steenbergen, [...], heeft een vriend als een kleine jongen laten misbruiken als opiumkoerier.
 ‘The great Rik van Steenbergen has let a friend be exploited as drugs courier.’

- (36) a. Zij lieten zich té gemakkelijk intimideren door de getruce Spaanse verdedigers. (EC: 2480)
 ‘They let themselves be intimidated too easy by the cunning Spanish defenders.’
- b. ?*Zij lieten de kinderen té gemakkelijk intimideren door de getruce Spaanse verdedigers
 ‘They let the children be intimidated too easy by the cunning Spanish defenders.’

Replacement of the affectee by a non-reflexive participant can sometimes lead to a grammatically well-formed sentence, but then we have to read the sentence as a (causeless) coercive CC: (34b) cannot be read as indicating permission, rather, it has to be read as denoting coercion. This implies a difference between the behavior of coercive reflexive constructions and permissive reflexive constructions: reflexive affectees of *coercive* CCs can be replaced.

The middle voice theory of Kemmer (1993) gives a good explanation for the relative difficulty in replacing the affectee in permissive reflexive constructions. Kemmer characterizes middle events as events in which there is *expected* coreference between the initiator and endpoint, or a *necessary* coreference between initiator and endpoint. According to Kemmer “the greater the participation or involvement of the endpoint self, the more likely the event is to be viewed as one that cannot be directed towards others” (p. 61 *ib.*). When Initiator and Endpoint must be the same entity, or when there is an expectation that they will be the same entity, *self-affectedness* is part of the meaning of the verb or construction (Kemmer 1993: 61).

So permissive reflexive CCs seem to denote middle events. Causer and affectee in reflexive permissive constructions have to be the same entity (the reflexive affectee cannot be replaced), so self-affectedness must be part of the meaning of this construction. Actions as denoted by the permissive reflexive CCs have become inherent self-directed actions, the kind of actions that cannot be directed towards others. Allowing someone else to be affected seems very strange. One has to have (some degree of) responsibility over the affectee, in order to be able to allow the affectedness of the affectee. This is closely connected with the pejorative meaning of most of these CCs. Allowing a negative effect is negative, but allowing a negative effect to happen to yourself is doubly negative. One has responsibility over oneself, and one should protect oneself. This

kind of negligence one can only have towards oneself (or towards affectees for which one is highly responsible), and this kind of events one can only “cause” at oneself.

The semantics of negated “middle” CCs, on the other hand, contain some positive judgement (see (37)).

- (37) Ik laat me niet intimideren!
 ‘I won't let myself be intimidated!’

Instead of letting oneself be affected negatively, one stood up for oneself.

As can be seen from examples (38) and (39), affectees from *coercive* reflexive constructions can be replaced.

- (38) a. Ze liet zich brengen naar gezinnen waar rouw heerste. (EC: 2755)
 She let herself bring to families were mourning was
 ‘She had herself taken to mourning families.’
 b. Ze liet haar zuster brengen naar gezinnen waar rouw heerste.
 ‘She had her sister taken to mourning families.’
- (40) a. Ook freule Poswick, de vroegere eigenaresse van de berg, heeft zich in houtskool laten vereeuwigen. (EC: 2733)
 Also freule Poswick, the former owner of the mountain has herself in charcoal let immortalize
 ‘Freule Poswick, the former owner of the mountain, also has had herself immortalized in charcoal.’
 b. Ook freule Poswick, de vroegere eigenaresse van de berg, heeft haar dochter in houtskool laten vereeuwigen.
 ‘Freule Poswick, the former owner of the mountain, also has had her daughter immortalized in charcoal.’

The explanation for this must lie in the fact that coercive reflexive constructions are no middle events and that self-affectedness (and expected coreference between initiator and endpoint) is not part of the meaning of these constructions. The distinction between middle events and reflexive events is crucial in Kemmer’s theory. Kemmer says that in reflexive events “the initiator acts on itself just as it would act on another entity; the reflexive marker is there simply to signal the unusual fact that the different participant roles happen to be filled by the same entity.” (p.

66 ib.). The type of coreference that exists between the causer and the affectee in reflexive coercive CCs is not different from the coreference in (40):

- (40) Ik zie mezelf in de spiegel.
 ‘I see myself in the mirror’

where initiator and endpoint do not have to refer to the same entity but can be replaced. This example and the coercive reflexive constructions can be interpreted as normal reflexive events.

Recall that the initial hypothesis of this paper predicts that permissive constructions have explicit causees. As we have seen, the corpus data did contain causeless permissive CCs, but they were all “middle” constructions; middle constructions are the only permissive constructions from the data with implicit causees. The reason for this deviant behavior of middle CCs might lie in the fact that focus is not on the activity of the causee, but on the negative self-affectedness of the causer. In this kind of events the causee has a marginal role, although he is autonomous. Not the causee is highly affected, but the causer is. The participant roles in the middle constructions are slightly different from the roles in normal permissive constructions. The responsibility of the causer for the realization of the effect is bigger, because he is also the (“participating”) endpoint in a negative event. The role of the causee automatically becomes smaller.

The identity of the causee can in some cases (instances of the middle CC) be irrelevant, and the focus can be on some property of the subject of the construction (as suggested by the ANS). However, this is not just any property, but some ‘bad’ property.

When the reflexive constructions are separated from the corpus data, the results from the corpus analysis give a different view. Table 2 shows the distributions in the different causeless CCs and the CCs with three participants in the corpus data, i.e. as in Table 1, but this time without the reflexive and middle CCs.

	<i>3 participants</i>		<i>Causeeless</i>		<i>Total</i>	
Permissive	7	8%	0	0%	7	3%
In-between	31	37%	39	33%	70	35%
Coercive	38	45%	73	62%	111	55%
Rest	8	10%	6	5%	14	7%
Total	84	100%	118	100%	202	100%

Table 2. *Distribution of inductive CCs (with laten) with three participants and without causee, without the reflexive and middle CCs*

When the reflexive and middle constructions are removed from the material, the corpus data confirm the expectation: the causeeless constructions are not found randomly in the material, the non-reflexive causeeless CCs from the corpus do not express permission.

3.4 Causeeless causative constructions with perception verbs in effected-predicates

Cognition and perception verbs as effects require actions from the causers of the construction, and not so much from the causees. The causees have a passive role, both in what they have to do to make the effected predicate happen, as in their wish for the effected-predicate to happen. Consequently, the causee does not seem to be an important participant in the conceptualization of the realization of the effected predicate. It is therefore not surprising that these constructions are well suited to leave the causee implicit: 41 out of the 72 constructions from the in-between group do not have an explicit causee. Some examples are:

- (41) Frankrijk heeft al laten weten, dat het tegen de uitbreiding geen bezwaren meer heeft [...]. (EC: 2449)
 France has already let know, that it against the expansion no objections anymore has
 ‘France already made known that it no longer objected to the expansion.’
- (42) De agent was kwaad, en liet dat duidelijk merken. (EC: 2734)
 The policeman was angry and let that clearly know
 ‘The policeman was angry, and made that very clear.’

- (43) Tien jaar geleden had ze een enorme hit met “the end of the world”, maar dat ze op die oude roem niet alleen hoeft te teren liet ze duidelijk horen en zien. (EC: 2400)

Ten years ago had she an enormous hit with “the end of the world, but that she on that old success not only has to live let she clearly hear and see

‘Ten years ago she had an enormous hit with “the end of the world”, but she clearly showed that she does not have to live on that success.’

Not only do constructions with perception or cognition verbs provide an opportunity for the causee to stay implicit, the constructions that actually do leave their causees implicit, get a somewhat separate meaning. Semantic properties that were already present in the instances with explicit causee have become stronger: the focus has moved more in the direction of the contact between the causer and the affectee, and has moved away from the causee, and from the actual realization of the effected-predicate. The focus is on the action of the causer. Whether there is actually someone who acts out the effected predicate, someone who “knows”, “sees” or “notices”, is unclear. Although experiencers are necessary participants in the conceptualization of cognition and perception verbs, in causative constructions with these verbs as effects, experiencers can be absent. Since we do not know if the effect is actually realized (and by who), it is not easy to add an implied causee. The examples can be best paraphrased as ‘to make known’ as in (41), ‘to make clear’ as in (42), and as ‘to show’ or even as ‘to prove’ (43).

So it seems that the constructions with cognition and perceptions verbs as their effected predicates have special characteristics (participant semantics) that allow them to leave the causee implicit. In turn, the constructions with implicit causees have developed their own semantics again.

3.5 Causelessness and the continuum of indirectness

Corpus analysis showed that the features coercive causation and permissive causation are not themselves sufficient conditions for leaving the causee implicit. As we have seen, all three distinct categories (permission, coercion and ‘in-between’) have possibilities for the causee to

stay implicit, but not to the same degree. It is some feature of the CC that allows the causee to stay implicit, and when the causee is indeed unexpressed, the causeeless CCs seem to develop their own characteristic meaning. Causeeless CCs in all three categories have their own semantics. In events expressed by coercive causeeless CCs, causees don't have to be forced, and they don't act against their wish: they seem to have some inherent tendency to carry out the effect. The goal of the causer is the effect only, the identity of the causee is irrelevant.

Causeeless CCs with perception or cognition verbs as their effected predicate do not need identifiable causees who actually carry out the effected predicates, and their effected predicates do not need to be realized. The focus is on some action of the causer, regardless of experiencers (causees). Permissive CCs show coreference between causer and affectee when the causee stays implicit. The causer refers to the same entity as the endpoint of the causal events, and its responsibility for the realization of the effect seems bigger than in normal permissive events, and the causee becomes less relevant.

How can these different CCs with and without causee be related to degree of indirectness (Kemmer & Verhagen 1994)? According to Kemmer and Verhagen:

“Directness vs indirectness of causation, we might notice, is closely related to the issue of which participant is more topical, the more indirect the causation, the more peripheral a participant the causee is, and peripherality normally would not cohere with topicality” (Kemmer & Verhagen 1994:133).

But the degree of indirectness of the CC does not seem to be directly related to (degree of) peripherality of the causee: the most indirect constructions (the ones with the most autonomous causees) are permissive constructions. Permissive constructions do not leave their causees easily implicit, as we should expect from constructions with peripheral causees. Coercive CCs on the other hand, easily leave their causees implicit, although these constructions are less indirect CCs, and their causees less autonomous.

However, the coercive CCs with implicit causees are not the most coercive constructions, in which the causees are forced to carry out the effect; causees that (for instance) have the effected predicates as their job have some inherent tendency to carry out the effect.

It might be the case that the constructions that leave their causees unexpressed can be found not at the ends of the continuum, but more towards the middle at both sides. Highly affected or integrated causees can be found at both edges of the continuum: at the least autonomous side, where the coercive constructions are placed, and at the most autonomous side, where the permissive constructions can be found. Here, causees are always expressed.

CCs that leave their causees implicit can be found more towards the middle of the continuum: these coercive constructions have more autonomous causees, and the permissive constructions less autonomous causees. Consequently, degree of autonomy cannot be related directly to degree of affectedness or peripherality. Highly integrated causees can be autonomous too, and causees of permissive constructions can be highly affected.

4. Conclusion

In this paper, I tried to find out which *laten*-CCs leave their causees implicit and which do not.

From Kemmer & Verhagen (1994) it could be concluded that implicit causees are extremely autonomous causees, and that explicit (zero marked) causees are the least autonomous causees. Least autonomous causees are part of a coercive construction, and highly autonomous causees are part of a permissive construction. Continuing this line of reasoning led to the expectation that causees of permissive CCs are most likely to be left implicit and causees of coercive CCs are least likely to be left implicit. The hypothesis of this paper was the opposite: causees of coercive constructions were most likely to be left implicit, and causees of permissive constructions are not left implicit, because they normally have a crucial role in a permissive (interpretation of the) event.

Data from the Eindhoven Corpus show that coercive constructions can easily leave their causees implicit, but coercive CCs that do, have their own specific semantics. Causees of these constructions have a strong tendency to carry out the effect, because it is their job or profession for instance. They do, however, have to get 'an order' from 'a client', who has the effect as his goal. Permissive constructions indeed do not leave their causees implicit easily, but the ones that do, do so under strict conditions, namely coreference between causer and affectee.

These “middle” CCs have their specific semantics too. Middle causative constructions express some pejorative sense, some negative value judgement towards the action of the subject of the construction. Initiator and endpoint (causer and affectee) refer to the same entity, and the effected predicate is most of the time negative. This means that the causer is negatively affected by an action he is the initiator of. The fact that the causer is “negatively self affected”, adds a negative value judgement to the content of the sentence: the subject should have acted differently, he should have had protected himself. Because self-affectedness is part of the semantics of the construction, the causer seems to get a bigger responsibility for the realization of the effect than in normal permissive events, and the role of the causee becomes smaller, or less autonomous.

CCs with perception or cognition verbs as their effected predicates can also leave their causees unexpressed, and get their own semantics when they have an implicit causee. The focus is on some action of the causer, an identifiable experiencer is no longer necessary in the conceptualization of the event. So, constructions from all three categories can have some special properties that allow the causee to be unexpressed, but the constructions that actually do have implicit causees seem to develop their own specific semantics and properties.

A related issue is how in general permissive or coercive meaning and the degree of autonomy of the causee have to be related to the degree of integration, affectedness and peripherality of the causee. According to Kemmer and Verhagen, coercive constructions have the most affected, integrated and topical causees, while permissive constructions have the least affected, integrated and topical causees. We have seen that this might not be the whole story. Zero marked causees are probably the most affected, integrated and focal causees (Kemmer & Verhagen 1994), but there seems to be no direct relation to autonomy of the causees, or to permission, or to indirect causation. Permissive causees of transitive causative constructions are zero marked too; after all, the constructions with ambiguous meanings (both permissive and coercive interpretations are possible) are the ones with zero marked causees. And so zero marked causees may very well be part of highly permissive constructions and thus be highly autonomous.

Coercive CCs that left their causees implicit were not the most coercive constructions, but the more moderate coercive ones. Consequently, the causees were not the least autonomous ones. On the conti-

num of indirectness, we would not put them at the end of one side, where we find the constructions with least autonomous causees, but more to the middle. Although it is probably the case that implicit causees can be characterized by a high degree of peripherality, and not being in focus, the relation with degree of autonomy cannot last. If there was a relation with autonomy, it would have been hardly possible to express some kind of coercive causation by causeless CCs, and causees of permissive constructions could not be highly affected or focal.

How then should we put the constructions on the continuum of indirectness, and how can we relate them to the different notions of affectedness autonomy and peripherality? On the continuum of indirectness, we find on one side the highly coercive constructions, and on the other side the permissive constructions, and in the middle the in-between-cases. The coercive constructions that leave the causees unexpressed are ‘moderately’ coercive (their causees are more autonomous), and can be found more in the middle of the continuum. The permissive constructions that leave their causees unexpressed, cannot be found at the end of the continuum either. Self-affectedness moves the focus (and affectedness) away from the causee, and the causee’s actions seem less autonomous. Consequently, they cannot be found at the end of the continuum either. Then we have the in-between CCs, which are located in the middle of the continuum anyway. It might be the case that the constructions with the most affected and integrated causees can be found at both ends of the continuum of indirectness, and that these values decrease when one goes to the middle of the continuum, by removing the causee, or by case marking of the causee. By placing highly affected causees on both sides of the continuum, the correlation with autonomy is no longer present. However, further research still has to be done on the conditions that determine affectedness of the causee, and that can give ultimate answers with respect to the relation between degree of affectedness of the causee and permissive/coercive causality.

References

- Algemene Nederlandse Spraakkunst*. 1994. Groningen: Wolters-Noordhoff.
- Boogaart, P.C. uit den (1975). *Woordfrequenties in geschreven en gesproken Nederlands*. Oosthoek, Scheltema and Holkema, Utrecht.
- Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. University of Chicago Press, Chicago.

- Degand, L. (2001). *Form and function of causation: A theoretical and empirical investigation of causal constructions in Dutch*. Peeters, Leuven.
- Dik, S.C. (1980). *Studies in functional grammar*. Academic Press, London/New York.
- Fagan, S.M.B. (1992). *The syntax and semantics of middle constructions. A study with special reference to German*. Cambridge University Press, Cambridge.
- Hopper, P.J. & S.A. Thompson (1980). Transitivity in Grammar and Discourse. *Language* 56:2, 251-299.
- Kemmer, S. (1993). *The middle voice*. John Benjamins, Amsterdam and Philadelphia.
- Kemmer, S. & A. Verhagen (1994). The grammar of causatives and the conceptual structure of events. *Cognitive Linguistics* 5, 115-156.
- Stukker, N., T. Sanders & A. Verhagen (1999). Waar een wil is is geen wet. De categorisering van causale relaties binnen en tussen zinnen. *Gramma/TTT, Tijdschrift voor taalwetenschap* 7, 66-86.
- Talmy, L. (1988). Force Dynamics in language and cognition. *Cognitive Science* 12, 49-100.
- Verhagen, A. & S. Kemmer (1992). Interactie en oorzakelijkheid. *Gramma/TTT, Tijdschrift voor taalwetenschap* 1, 1-20.
- Verhagen, A. & S. Kemmer (1997). Interaction and causation: A cognitive approach to causative constructions in Modern Standard Dutch. *Journal of pragmatics* 27, 61-82.
- Verhagen, A. (2000). Interpreting Usage: Construing the history of Dutch causal verbs. In M. Barlow & S. Kemmer (eds.), *Usage based models of language*, 261-286. Stanford: CSLI Publications.
- Waltereit, R. (2000). What it means to deceive yourself: The semantic relation of French reflexive verbs and their corresponding transitive verbs. In Z. Frajzyngier & T.S. Curl (eds.), *Reflexives. Form and function*, 257-278. Benjamins, Amsterdam.

Drift in Dutch: Fleshing out the factors of change

Thomas F. Shannon (University of California at Berkeley)

1. Introduction

A well-known adage in linguistics has it that all languages change over time. The Dutch language is certainly no exception to this rule, whether over the full course of its documented history or even in the relatively short space of just a century. For instance, anyone who looks at texts from the late 19th century cannot help being struck by differences in the language of that day vis-à-vis its 21st century descendant. Let us take as an example a brief passage from the novel *Uit het Leven* by Arie Prins (1885: 4).¹

De **groot**e schorsteenen wierpen dikke rookwolken uit, waarin **ee**ne menigte vonken dwarrelden, en Spinoza hoorde de regelmatige, heldere slagen **der** hamers op het ijzer, het sissend aflaten van **den** stoom en het dof gekraak **der** machines, die in beweging waren. ...

Haastiglijk liep hij vervolgens **tusschen** **ee**nige verroeste, oude stoomketels door, de morsige werkplaats over en schoof de **groot**e draaijerij in. In deze werkplaats, – een groot vertrek, met gewitte muren en kleine boogramen, waar een grijs daglicht door viel, dat aan alles, niet tegenstaande de bedrijvigheid, een koud, dood**sch** aanzien gaf, – maakten de raderen van een tiental ijzeren draaibanken een **helsch** leven.

(‘The big chimneys threw thick smoke clouds out, in which a bunch of sparks swirled, and Spinoza heard the regular, clear blows of the hammers on the iron, the hissing release of the steam and the dull clank of the machines that were in action. ...

Hastily he then ran between some rusty, old steam kettles, over the dirty workplace and pushed on into the large turning mill. In this work area – a large room, with whitewashed walls and small arched windows, through which a gray light fell that gave everything, despite the activity, a cold, dead appearance – the wheels of ten iron lathes created a hellish life.’)

¹ Forms to pay attention to in this passage and below are highlighted in **bold**.

In this short excerpt several points jump out at us where this text differs linguistically from its modern counterparts.² For instance, changes in spelling make such earlier texts appear unusual to 21st century readers. Thus, in modern Dutch words like the following are now spelled with a single rather than double vowel in open syllable to indicate vowel length:

- (1) a. *groot* ‘big’, *rood* ‘red’, *boomen* ‘trees’, *zoo* ‘so’, *oogen* ‘eyes’, *loopen* ‘run’
 b. *eene* ‘a’, *eenige* ‘some’, *beenen* ‘legs’, *meening* ‘opinion’, *bleeke* ‘pale’

Moreover, modern Dutch has replaced the cluster *sch* by *s* in non-initial position (except in proper names like the Dutch beer *Grolsch*):

- (2) *tusschen* ‘between’, *doodsch* ‘dead’, *helsch* ‘hellish’, *asch* ‘ash’, *flesch* ‘bottle’

Furthermore, the elision of intervocalic *d* in many words is regularly represented in modern Dutch orthography, whereas a century ago the full, unelided form was still found very frequently in texts (though not in the passage cited above):

- (3) *neder* ‘down’, *weder* ‘again’, *medenemen* ‘take along’ > *neer*, *weer*, *meenemen*

However, the differences noted so far are for the most part rather superficial and arguably largely effect the external appearance of the written language. But a corollary to the adage about continual change in language alluded to above has it that languages change in all aspects, particularly in all facets of their grammar. Accordingly, other changes that we can also note in this work are more substantial in apparently affecting the very grammar of the language. For example, written texts from the late 19th century seem to preserve more inflectional morphology, particularly in noun phrases, than those from the 21st century. Thus,

² In the following, the examples are taken from the first ten pages of *Uit het Leven*, which include the passage cited above. Here and later, “modern” designates the 20th and 21st centuries.

in instances such as (4) the indefinite article and possessive adjectives regularly showed the agreement suffix *-e* in feminine and plural forms of determiners, which is no longer found in modern Dutch.

- (4) *eene menigte* ‘a bunch’, *zijne zware stem* ‘his heavy voice’, *hare oogen* ‘her eyes’; *hunne lippen* ‘their lips’, *hunne hoofden* ‘their heads’, *zijne smerige handen* ‘his dirty hands’

Furthermore, nominal case marking appears to have been better preserved over a century ago, at least in writing. Thus we find oblique marking on noun objects of prepositions in (5), and genitive marking on (feminine singular or plural) possessive nouns in (6). Moreover we also find oblique (accusative/dative) marking on masculine singular nouns as verbal direct (7a–b) and indirect (7c) objects. Finally, the example in (8) even shows a subjunctive form of the verb *zijn* ‘to be’.

- (5) *aan den wand* ‘at the wall’, *in den mist* ‘in the mist’, *in vollen gang* ‘in full gear’, *van den stoom* ‘of the steam’, *door den baas* ‘by the boss’, *op den grond* ‘on the ground’
- (6) *het dof gekraak der machines* ‘the dull clank of the machines’, *de ... slagen der hamers* ‘the ... blows of the hammers’, *op den hoek der straat* ‘on the corner of the street’
- (7) a. *Moet ik den meester gaan halen?*
‘Must I go get the master?’
b. ... *maar hij voltooide den zin niet.*
‘... but he didn’t finish the sentence.’
c. ... *hoeveel fooi hij den debiteur gegeven had ...*
‘... how much tip he had given to the bookie ...’
- (8) ... *indien het ruchtbaar ware geworden, dat hij dronk.*
‘... if it had become known that he drank.’

Now these examples demonstrate for the most part commonplace instances of morphological simplification. But is it also possible to observe similarly striking differences in the syntax of the language of such late 19th century Dutch texts, for example in word order? Indeed, it is often felt that since such syntactic change takes place over very long

stretches of time it is much less easily observed in the space of a mere century. However, in reviewing a number of possible changes in Dutch over the past hundred years, Joop van der Horst (1995: 70) briefly remarks:

Wie nu teksten van rond de eeuwwisseling leest, kan bijna van zin tot zin de verschillen aanwijzen met het hedendaagse Nederlands. Een opvallend verschil is de positie van voornaamwoorden: die stonden omstreeks 1900 nog dikwijls vóór het onderwerp van de zin: *Als hem de rector van school stuurt*. Destijds was heel gewoon: *Hoe maken het je zootjes?* Misschien is deze volgorde nog steeds mogelijk, maar zulke zinnen klinken nu wel ouderwets.

(‘Whoever now reads texts from around the turn of the century can almost from sentence to sentence point out the differences from contemporary Dutch. One striking difference is the position of pronouns: around 1900 they still often stood before the subject of the clause: *Als hem de rector van school stuurt*. [‘If the rector sends him from school’]. At that time [the following] was very common: *Hoe maken het je zootjes?* [‘How do it your sons’, i.e. ‘How are your sons doing?’] Perhaps this word order is still possible, but such sentences now sound very old-fashioned.’)

It is not apparent what specific types of examples Van der Horst had in mind here. Surely it is not being claimed that pronoun objects can never be found before a medial subject noun in modern Dutch, but Van der Horst does not indicate what it is about these sentences that makes them exceptional. Nevertheless, his general claim is quite patent. He is apparently suggesting that there has been a shift in Dutch away from placing medial noun subjects after pronoun objects in the course of the last century, although he does not give much evidence to support this claim. In the present study I would like to do just that: I will examine empirically the question whether there has been such a change in Dutch medial word order over the past hundred years. Furthermore, we wish to place these developments within a larger historical context, for it will be seen that this question has a longer history and in fact must be viewed against the background of other developments in the language, at least some of which were alluded to above. Below we will offer evidence that there has indeed been such a change in Dutch, one which has been going on for centuries and apparently is approaching its culmination only now, in the 21st century.

2. The order of Dutch subjects and objects over the centuries

This paper examines the linear ordering of (pro-)nominal constituents in the so-called “middle field” (i.e. within the sentence brace) in the Dutch language, specifically the order of pronominal objects and nominal subjects there.³ In previous work (Shannon 1997) I briefly examined a claim made by van Gestel et al. (1992) that while in Middle Dutch, pronoun objects were regularly placed before noun subjects, in modern Dutch that is no longer possible. To check this, I studied the order of these elements in the 16th century *Ulenspieghel* and its modern Dutch translation. This comparison furnished very good prima facie evidence that the order pronominal object before nominal subject – which I refer to for convenience as “(pronoun) object preposing,” without attaching any theoretical significance to the movement metaphor invoked by the term – was indeed the dominant order in the Middle Dutch text, whereas the modern Dutch translation overwhelmingly evinced the opposite order (dubbed “object postposing”, again for convenience).⁴ Subsequently, these orderings in modern Dutch and German were studied in Shannon (2000), where it was found that in modern Dutch prose fiction texts pronoun object postposing clearly dominates, whereas in such modern German texts pronoun object preposing is the rule.

However, this previous work was somewhat limited in scope and coverage and the present study attempts to address this shortcoming. For this investigation, I collected and compared Dutch corpora covering the 16th, 19th, and 20th centuries. These corpora were comprised of the following works (sources and abbreviated reference forms are given at the end of this study):

³ On the positional fields model adopted here, see Haeseryn et al. (1997: 1225–34). Briefly, the middle field (*het middenstuk*) is the sentence area between the first prong or pole (“complementizer” or C in generative terms), which is formed in main clauses by the finite verb in first or second position, in subordinate clauses by the subordinating element, and the second prong, formed by the nonfinite verbal elements at the end of the clause. Most clausal elements are usually contained within the middle field.

⁴ Similarly, when speaking of the position of the noun I will refer to subject pre- vs. postposing.

- 16th century: Geeraedts (1986, orig. ca. 1535), Lecoutere & de Vreese (1904, orig. 1554), Schellaert (1952, orig. 1516), de Vreese & de Vries (1941, orig. 1501).
- 19th century: Buysse (1893), Daum (1889), Heijermans (1893), Prins (1885).
- 20th century: Claus (1989), Gijzen (1980; first 150 pages), Van der Heijden (1992), De Moor (1993), Mulisch (1987), Nooteboom (1991).

These texts were selected because they came from three representative and rather well-documented centuries in the historical development of the Dutch language and were prose fiction works, in which considerations of rhyme or meter, scholarly style, etc. would presumably exert little or no influence.⁵ The 16th century texts are popular prose works from the first half of the century and were selected primarily for reasons of availability. The 19th and 20th century texts were written approximately 100 years apart (between 1880–1895 and 1980–1995, respectively) by recognized authors who presumably reflect the syntactic norms of their time.

Although other data for the 16th and 20th centuries have been reported previously (Shannon 1997, 2000), the corpora for this study are all different and/or expanded. For the 16th century this was particularly necessary because the earlier work cited only limited data from a single text, the *Ulenspieghel*, which is quite likely a translation of a lost Low German text. For the present study, three other chapbooks (*volksboeken*) were added. Moreover, although the 20th century Dutch results reported in Shannon (2000) were quite extensive, all new texts were consulted for the current study. Finally, since we wanted to empirically test Van der Horst's claim about the difference between 19th and 20th century Dutch word order, we added a corpus of works from the 19th century; these data are reported here for the first time. It should be noted that there is a size

⁵ While every effort has been made to obtain sufficient data to warrant generalizations, our corpora still represent only a small sample of the range of texts produced at any time. Consequently, the results reported here can only claim limited reliability – for certain (kinds of) written texts. However, given the clear quantitative data reported here, it is quite likely that our observations do in fact represent actual developments in the Dutch language as a whole.

difference here: the 16th and 19th century corpora each contain four works, the modern Dutch corpus six. This is because the modern Dutch works tended to be shorter, and so to obtain more data additional modern Dutch novels were examined (from the one longer text – Gijzen 1980 – only the first 150 pages were excerpted). Nonetheless, the 19th century and 20th century corpora are roughly of equal length in terms of total words, while the 16th century corpus is somewhat shorter: 16c = 146,500, 19c = 189,200, 20c = 194,000.⁶ In comparisons below the size difference should be kept in mind. Overall, the results presented here offer a much broader view of the evolution of pronoun object preposing in Dutch than the earlier studies.

In each work all examples with a nominal subject (i.e. one containing a full noun head) and pronominal object (one with a pronoun head) in the middle field were collected.⁷ Only examples containing a full finite clause with at least the first prong present were included. Thus, elliptical clauses such as (9) and nonfinite constructions with participles

⁶ In figures and examples the century is abbreviated, e.g. 16c = 16th century. To estimate word lengths, the average number of words was calculated for five random pages from each work and multiplied by the number of pages. The approximated word counts were as follows. 16c: *Ulenspieghel* 23,000; *Margarietha van Lymborch* 94,500; *Turias ende Floreta* 20,800; *Salomon ende Marcolphus* 8,200. 19c: Buysse 48,300; Daum 71,200; Heijermans 39,500; Prins 30,200. 20c: Claus 22,000; Gijzen 56,500; van der Heijden 21,200; de Moor 46,100; Mulisch 20,100; Nooteboom 28,100.

⁷ It must be stressed that we are discussing the distribution of phrasal, not lexical categories, i.e. the subjects and objects in question consist of noun phrases headed by nouns and pronouns, respectively. For convenience we will talk loosely of “noun/nominal subjects” and “pronoun/pronominal objects”, which is perhaps misleading in that it seems to suggest that these elements comprise just a single word. Though pronouns do tend to be one word long, they can be longer, as when they are modified; moreover, writing conventions as to what constitutes a single word can change. Some examples from our corpora are: *hem lieden* ‘them folks’, *haer selven* ‘her self’; *zich zelf* ‘one self’, *hen beiden* ‘them both’, *zelfhaar* ‘even her’. Similarly, although proper nouns often consist of one word, this is not always the case; cf. the following examples from our database: *den Ekster*, *oom Robert* ‘Uncle Robert’, *prof. Felice* ‘Professor Felice’, *Mevrouw Sorgeloos die gescheiden is* ‘Mrs. Sorgeloos, who is divorced’. So, despite the labels “noun/pronoun object”, it must be kept in mind throughout that we are actually referring to phrases, not just words.

such as (10), as well as verb-late main clauses like (11) and clauses where one element was in the middle field but the other was exbraciated to the postfield (12) were excluded.⁸ All such examples were found only in the 16th century corpus, and even there they were rare.

- (9) 16c Etsijtes die gheraecte aen den coninc van tartarijen die hi ooc dapperlic aen ghinck. Ende *die coninc* **hem** weder dier gelike. [S 118]
 ‘Etsijtes, he came upon the king of Tartaria, whom he bravely attacked. And the king [did] the same to him again.’
- (10) 16c *Turias* **dat** hoorende sprack totten schiltknecht ... [L 19]
 ‘Turias hearing that spoke to the squire ...’
- (11) 16c Noyt *scheyden* **mi** so *deerde* alst dede dat sceiden vander schoonder maecht. [S 16]
 ‘Never parting so hurt me than (it) did the parting from the beautiful maiden.’
- (12) 16c so comt ... te constantinobele daer wert **u** *gheclaecht alle mijnen noot*... [S 52]
 ‘So come ... to Constantinople where to you will be complained all my need ...’

All examples found were entered into a Macintosh PowerBook, using the database program Panorama, and tagged for factors which were hypothesized to be correlated with word order.⁹ Where appropriate, chi-square tests were run using Microsoft Excel.¹⁰ The following sections present the results of that analysis, first giving the overall findings (§2.1),

⁸ For ease of reference, noun subjects are in *italics*, pronoun objects in **bold**. At times **bold italics** are also used in order to distinguish other items (e.g. 1st or 2nd prong, indirect vs. direct object).

⁹ One potentially important factor, theme-rheme (cf. also Behaghel’s second law) or information structure, was not considered here because it proved too difficult to evaluate it reliably.

¹⁰ Robert Kirsner (p.c.) has pointed out that the traditional use of chi-square in corpus-based, empirical linguistic studies has recently been questioned. We continue to use it here with caution, however.

then examining more specific correlations between properties of the elements and the word order found (§2.2).

2.1. Overall results

First we report the aggregate results for all examples found, including all types of pronominal objects in our three corpora.¹¹ The figures are given in table 1 below.

	16 th century	19 th century	20 th century
NSubj + ProObj	27% (158)	74.4% (300)	87.5% (265)
ProObj + Nsubj	73% (428)	25.6% (103)	12.5% (38)
Total	100% (586)	100% (403)	100% (303)

Table 1: *Order of Dutch noun subject and pronoun object for three centuries, all examples*

These data clearly document a continuing and ultimately drastic syntactic shift in linearization over the past 500 years in Dutch, as well as providing prima facie evidence supporting Van der Horst's contention that there has been a change in medial word order in the course of the last century. Pronoun object preposing has dropped from around 75% in the 16th century¹² to 25% in the late 19th century – the figures for these two centuries are thus almost exactly reversed! – and then to 12% in the late 20th century texts. These differences are highly significant, according to the chi-square test ($p < .001$). We seem to be following a continuing transformation of Dutch word order from the Middle Ages through the 20th century and beyond. This change appears to be still underway, but the general direction is evident: a clear shift from pronoun object preposing to postposing. In fact, the older preferred order is at present

¹¹ Compound pronouns like *hem/hen lieden* 'them folks', which were found in the 16th century, were included here, although they were quite infrequent. Similarly, compound reflexives like *zich zelf* (now written as one word) were also counted.

¹² Moreover, the 16th century corpus contains a high percentage of postposed demonstrative pronouns. If we discount demonstratives, the frequency of object preposing data is even greater: cf. table 3 below. Percentages cited in the text are rounded off, those in tables are given to one or two decimal places.

rarely found and may soon vanish – as has long since been the case in English.

2.2. Apparent relevant factors

From more detailed analysis of the data, we can establish that certain factors are related to the linearization of these elements. Below we examine such properties, first for the pronoun object (2.2.1), and then for the nominal subject (2.2.2).

2.2.1. Linear order and pronoun object properties

Let us first consider possible correlations between pronoun object properties and linearization. A number of pronoun properties were tracked, which will be reviewed here: pronoun type, case, animacy, and clitic (reduced) status.

2.2.1.1. Pronoun type. Several types of pronouns were distinguished: personal, reflexive, demonstrative, reciprocal (as far as it was distinct from the reflexive), and indefinite. Table 2 gives the breakdown of object preposing by pronoun object type in our three corpora.¹³

	16 th century	19 th century	20 th century
Personal			
NSubj + ProObj	12.8% (43)	76.5% (192)	89.8% (141)
ProObj + NSubj	88.2% (294)	23.5% (59)	10.2% (16)
Subtotal	100% (337)	100% (251)	100% (157)
Reflexive			
NSubj + ProObj	30.0% (18)	63.6% (77)	81.0% (94)
ProObj + NSubj	70.0% (42)	36.4% (44)	19.0% (22)
Subtotal	100% (60)	100% (121)	100% (116)

¹³ In our 16th century corpus the true reflexive pronoun *zich*, a borrowing from German, does not yet occur and forms of the anaphoric (personal) pronouns are instead used in reflexive function. The instances of reflexives in the 16th century texts were therefore determined solely by meaning and not by form. The later centuries show exclusive use of the reflexive *zich* (or *zichzelf*) in this function.

Demonstrative			
NSubj + ProObj	51.1% (96)	100% (6)	100% (11)
ProObj + NSubj	48.9% (92)	0% (0)	0% (0)
Subtotal	100% (188)	100% (6)	100% (11)
Reciprocal			
NSubj + ProObj	100% (1)	100% (2)	100% (6)
ProObj + NSubj	0.0% (0)	0.0% (0)	0.0% (0)
Subtotal	100% (1)	100% (2)	100% (6)
Indefinite			
NSubj + ProObj		100% (23)	100% (13)
ProObj + NSubj		0% (0)	0% (0)
Subtotal		100% (23)	100% (13)
Total	100% (586)	100% (403)	100% (303)

Table 2: *Order of Dutch noun subject and pronoun object for three centuries as a function of pronoun type*

As these figures clearly show, various types of pronoun objects are differentially prone to preposing. Personal and reflexive pronouns most heavily favor pronoun object preposing, whereas demonstratives even in Middle Dutch only preposed about half the time, and neither indefinites nor reciprocals appear to have had any proclivity for preposing (admittedly, the data for reciprocals is scant, but their penchant nonetheless appears quite certain). While the reason for this discrepancy is not completely clear, undoubtedly it is not due to differences in length, as pronouns tend overwhelmingly to consist of a single word (cf. §2.2.3). Presumably this varying behavior is due to other factors such as the differing semantics and/or pragmatics of these pronoun types, or, as we will suggest (§4.3), their morphological distinctness. Whereas personal and reflexive pronouns tend to refer to previously established, often more backgrounded discourse referents, demonstratives – and perhaps reciprocals – tend to be less backgrounded and possibly stressed, and indefinites normally do not designate already known discourse entities. The former were also in the 16th century more morphologically distinct. Furthermore, indefinite subjects also evince a propensity to appear late in the middle field in our data (§2.2.2.2).

These results largely agree with those reported in Shannon (1997, 2000). In particular, for those pronouns which early on show a tendency to prepose (especially personal and reflexive pronouns, but also – at least

initially – demonstratives) there is a constant, significant trend toward less frequent preposing over time. Interestingly, although in the 16th century personal pronoun objects are preposed more often than reflexives (almost 90% for personal pronouns, as opposed to 70% for reflexives), by the 19th and 20th centuries pronoun object preposing had declined more rapidly with personal pronouns, so that reflexives now prepose more frequently than personal pronouns (approximately 36% vs. 24% in the 19th century, and 19% vs. 10% in the 20th). Chi-square indicates that these differences between personal and reflexive pronouns are significant ($p < .0001$, .01, .04, respectively). While it is not certain why reflexive pronouns have been so much more resistant to the drift away from pronoun object postposing, it may have to do with the type of predicate that reflexives are often associated with (cf. §2.2.2.4 on the semantic role of the subject). In addition, their unambiguous marking as objects has also helped reflexives resist the drift toward postposing (cf. §4.3).

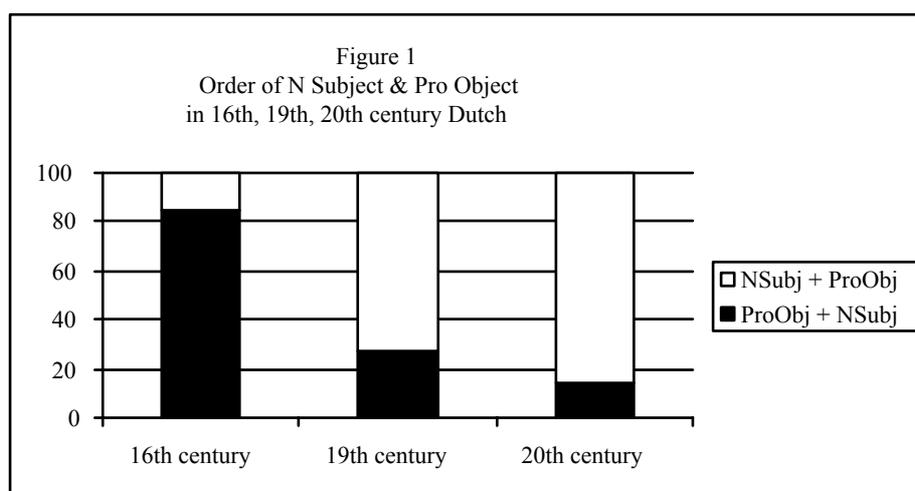
Recall that our previous figures for pronoun object preposing in table 1 included all pronoun types. Since only personal and reflexive pronouns have ever showed any real disposition to prepose, it is instructive to exclude other types in studying pronoun object preposing. The following table gives the figures on object preposing for personal and reflexive pronouns only.

	16 th c.	19 th c.	20 th c.
NSubj + ProObj	15.4% (61)	72.3% (269)	86.1% (235)
ProObj + NSubj	84.6% (336)	27.7% (103)	13.9% (38)
Total	100% (397)	100% (372)	100% (273)

Table 3: *Order of Dutch noun subject and pronoun object for three centuries, only personal and reflexive pronoun objects*

It appears that the frequency of cooccurrence of nominal subjects and pronominal objects in the middle field has decreased over time in Dutch. The frequency of examples shows a decline, from 2.7 occurrences per 1,000 words in the 16th century data (397 occurrences, ca. 146,500 words), to 2.1 occurrences in the 19th century corpus (372 occurrences, ca. 189,200 words), and 1.4 in the 20th century (273 occurrences, ca. 194,000 words). More importantly for our immediate concerns, we find here that the shift in frequencies of object preposing is even more dramatic than table 1 had indicated, going now from 85% in the 16th century to 28% in the 19th century and falling finally to a low of 14% in

our 20th century corpus. Even without statistical analysis, these differences are clearly substantial, but chi-square confirms that they are highly significant. Figure 1 below gives the same data for these three periods in the form of a bar chart which graphically displays the change over time.



Since only personal and reflexive pronoun objects have favored preposing before medial subjects in the periods covered, we will cite figures on these two pronoun types alone in the rest of this study. Below we give a few representative examples with demonstratives and indefinites.

- (13) a. 16c Als **dat** *die grave* hoorde soe vraechde hi wat die bede was. [S 34]
 ‘When the count heard that he asked what the request was.’
 b. 16c Als *die coninc* **dat** hoorde, doen was hy seer blijde ... [S 16]
 ‘When the king heard that he was very glad ...’
- (14) a. 19c Weet *je moeder* **dat**? [Hm 32]
 ‘Does your mother know that?’

- b. 19c Zij voerde een stillen oorlog, waarvan *haar man niets* merkte, ... [P 93]
 ‘She waged a quiet war, of which her husband noticed nothing.’

- (15) a. 20c Wist *mijn Mama dat*? [C 59]
 ‘Did my mother know that?’
 b. 20c Richard, ... moeten *de buren alles* horen? [C 92]
 ‘Richard, ... must the neighbors hear everything?’

2.2.1.2. Pronoun case. Next we examine the possible correlation of pronoun case with (personal and reflexive) pronoun preposing; only dative and accusative case are considered.¹⁴ The overall results are presented in the following table.

	16 th c.	19 th c.	20 th c.
Dative Pronoun			
NSubj + ProObj	13.6% (21)	73.7% (73)	86.9% (53)
ProObj + NSubj	86.4% (133)	26.3% (26)	13.1% (8)
Subtotal	100% (154)	100% (99)	100% (61)
Accusative Pronoun			
NSubj + ProObj	15.9% (38)	71.8% (196)	85.8% (182)
ProObj + NSubj	84.1% (201)	28.2% (77)	14.2% (30)
Subtotal	100% (239)	100% (273)	100% (212)
Total	100% (393)	100% (372)	100% (273)

Table 4a: *Order of Dutch noun subject and pronoun object for three centuries based on pronoun case (personal and reflexive pronoun objects)*

¹⁴ In addition, the 16th century corpus contained four genitive pronouns, two preposed, two postposed, which are not tallied in the table. Since Dutch does not display a clear morphological distinction between oblique forms, case was assigned based on what it would presumably have been if there were clear morphological distinctions (e.g. as would be found in German). In order to determine whether there is a correlation between case and linearization it was of course crucial to make this distinction. While it is arguably risky to assign case in the absence of clear morphological markers, it does not appear that many of the assignments would be controversial. Most of the datives correspond to traditional types of “indirect object” such as second object of a verb of transfer, dative of possession, certain experiencers, etc.

Viewed in this global fashion, no difference is found with respect to preposing as a function of case: in each century, datives and accusatives are preposed with about the same frequency. The gap is only about one or two percentage points, with the accusative enjoying the slightly higher frequency in all instances; chi-square reveals no significant differences. However, if we consider the type of pronoun involved, things look slightly different, as we see in the next table.

	16 th c.	19 th c.	20 th c.
Personal Pro. Dat.			
NSubj + ProObj	13.6% (21)	73.1% (68)	85.5% (47)
ProObj + NSubj	86.4% (133)	26.9% (25)	14.5% (8)
Subtotal	100% (154)	100% (93)	100% (55)
Personal Pro. Acc.			
NSubj + ProObj	11.2% (20)	78.5% (124)	92.2% (94)
ProObj + NSubj	88.8% (159)	21.5% (34)	7.8% (8)
Subtotal	100% (179)	100% (158)	100% (102)
Subtotal Personal Pro.	100% (333)	100% (251)	100% (157)
Reflexive Pro. Dat.			
NSubj + ProObj	(0)	83.3% (5)	100% (6)
ProObj + NSubj	(0)	16.7% (1)	(0)
Subtotal		100% (6)	100% (6)
Reflexive Pro. Acc.			
NSubj + ProObj	30.0% (18)	62.6% (72)	80.0% (88)
ProObj + NSubj	70.0% (42)	37.4% (43)	20.0% (22)
Subtotal	100% (60)	100% (115)	100% (110)
Subtotal Reflexive Pro.	100% (60)	100% (121)	100% (116)
Total Both Types	100% (393)	100% (372)	100% (273)

Table 4b: *Order of Dutch noun subject and pronoun object for three centuries based on pronoun case and type (personal and reflexive pronoun objects)*

These results are quite comparable with those reported in Shannon (2000: 156ff.). Looking first at personal pronouns, we find that although preposing has declined for both cases, the decline has been less rapid with datives than with accusatives. While in the 16th century corpus both cases evince preposing with almost equal frequency (dative 86%, accusative 89%), in the 19th century preposing drops somewhat lower for accusatives (89% > 22%) than for datives (86% > 27%), and in the 20th century only dative forms prepose with a frequency greater than 10%

(dative 15%, accusative 8%). It appears that there has been a somewhat greater resistance to the shift toward postposing among datives than among accusatives. According to chi-square these differences are not statistically significant, however, even for the 20th century (16c $p < .6$, 19c $p < .4$, 20c $p < .2$), although the correlation is obviously becoming stronger.

However, despite the fact that chi-square turned up no significant correlation with case, several Dutch colleagues have said they find preposing a pronoun more acceptable when it is an indirect object, so it may be true that the difference in case is related (sometimes) to linearization.¹⁵ It is not clear why these differences based on case, small as they are, exist.¹⁶ One possibility is that the dative personal pronouns often represent human experiencers, whereas the preposed accusative pronouns may represent (inanimate) patients. Presumably the human experiencers are more often topical elements, as human experiencers appear to be inherently more contextually given (cf. the next section). If so, then the

¹⁵ I am indebted to Arie Verhagen (p.c.) for reminding me that there is in fact something of a tradition to the claim that in Dutch (virtually) only indirect objects (called “datives” here), in fact only of verbs which select *zijn* ‘be’ as their perfect auxiliary such as *bevallen* ‘to please’, allow pronoun object preposing. Cf. e.g. Koster (1978), Balk-Smit Duyzendkunst (1979), Verhagen (1986: ch. 6). As Balk, Verhagen, and the present study demonstrate, however, this claim is not factually correct: although such verbs do seem to favor object preposing, they are not the only ones which do.

¹⁶ Verhagen (1986: ch. 6) proposes an interesting functional account of Dutch word order which potentially accounts for these differences. Verhagen argues that indirect objects (datives) are always independent participants (not fully affected), whereas direct objects (accusatives) are only sometimes independent. He ties this into his more general claim that the “order of two NPs implies that at least the first of them is perceivable independently of the evoked state of affairs” (p. 225). In this account, the greater proclivity of indirect objects as opposed to direct objects to prepose would then follow from the former’s status as more often independently perceivable.

While Verhagen’s arguments for this analysis of modern Dutch are on the whole convincing, as it stands the analysis cannot be fully correct, since it does not explain how it is that pronoun object preposing has shifted in Dutch over time. In its absolute form, Verhagen’s account would lead us to expect that instead of changing these patterns should have remained constant.

disparity may be attributable to the difference in topicality and/or animacy and less to case itself. However, it may also be that the effect is due to the semantic role of the accompanying subject (cf. §2.2.2.4), as dative experiencers are often found with nonagentive (usually inanimate) subjects (cf. 25b later). And with such asymmetrical semantics, ambiguity as to subject and object would not arise (cf. §4.3).

For reflexives, the data on datives is rather scant, but what we do find suggests that here the situation is reversed, namely dative reflexives prepose less frequently than accusatives (though there are not enough cases to warrant applying chi-square). Of the twelve cases of dative reflexives in our data, only one is preposed (19th century), whereas with accusative reflexives preposing continues to decline less rapidly than it does with personal pronouns in either case: 16c 70% (42/60); 19c 37% (43/115); 20c 20% (22/110).

Note finally that the overall frequency of object pronoun forms actually declines over time, but differentially. While the frequency of reflexive forms per 1,000 words has remained relatively constant (accusative: 16c = .41, 19c = .61, 20c = .57; we disregard dative reflexives as too rare), the frequency of personal pronouns has declined substantially, especially for the datives (dative: 16c = 1.05, 19c = .49, 20c = .28; accusative: 16c = 1.22, 19c = .83, 20c = .53). Given that personal pronoun objects initially displayed the highest proclivity to prepose and that the dative personal pronoun shows a slightly greater frequency of preposing in the 19th and 20th centuries, it would seem likely that the marked decrease in the frequency of personal pronouns, especially datives, is related to the overall decrease in the frequency of object preposing.

2.2.1.3. Pronoun animacy. Next we examine the correlation between preposing and the animacy of the pronoun object. Here we use a simple dichotomy of animate (human and animate in the strict sense) versus inanimate (concrete and abstract are conflated under this rubric). Since true reflexive pronouns do not have inherent animacy but inherit it from their subject, reflexives are not counted in the tally, but rather only personal pronouns. Moreover, even though in the 16th century personal pronouns were used in reflexive function, for the sake of comparison we also eliminate those cases here and consider only personal pronouns used anaphorically. The results are listed in table 5. All examples involving a

personal pronoun are considered, including the four instances of genitive objects (two preposed, two postposed).

	16 th c.	19 th c.	20 th c.
Personal Pro. Animate			
NSubj + ProObj	14.2% (43)	72.3% (149)	87.8% (115)
ProObj + NSubj	85.8% (260)	27.7% (57)	12.2% (16)
Subtotal	100% (303)	100% (206)	100% (131)
Personal Pro. Inanimate			
NSubj + ProObj	(0)	95.6% (43)	100% (26)
ProObj + NSubj	100% (34)	4.4% (2)	(0)
Subtotal	100% (34)	100% (45)	100% (26)
Total	100% (337)	100% (251)	100% (157)

Table 5: *Order of Dutch noun subject and pronoun object for three centuries based on pronoun animacy (personal pronoun objects only)*

Except for the somewhat anomalous data for the 16th century, the general trend is clearly for animate pronoun objects to prepose more often than inanimates. Actually, the data for the 16th century are too scant to offer a fruitful comparison – and preposing was the rule anyway; a chi-square test does not reveal a statistically significant distinction here. However, for the 19th century results chi-square does show significance, $p \leq 0.001$; and the 20th century data just miss significance at the 0.05 level. Once again, these results agree well with those of Shannon (2000: 153ff.), where it was noted that with personal pronouns animates tend to prepose more than inanimates do. Presumably this skewing is due to an anthropocentric bias, such that animate entities (particularly humans) are inherently of more interest and hence tend to be mentioned earlier than inanimates. But compare also our later remarks on subject semantic role (2.2.2.4).

2.2.1.4. Pronoun cliticization. Finally, let us examine the possible relation between clitic status of the pronoun (cf. Booij 1995: ch. 8 on modern Dutch) and linear order. Clitic status is of potential relevance here because according to Wackernagel's law (cf. e.g. Lehmann 1992: 285) such weak forms are held to attract – in the older languages at least – to clause-second position, viz. the first prong in the positional fields model (C in generative analyses), and hence should precede nonclitic

elements such as nominal subjects. Below we crosstabulate the findings for our three corpora based on the clitic status of the pronoun object; for comparison, the relevant data are added from the database on 20th century Dutch gathered for Shannon (2000). In these results we follow the traditional definition of clitics: all pronoun objects whose written shape clearly indicated a phonologically reduced form were considered to be clitics.¹⁷

	16 th c.	19 th c.	20 th c.	Shannon (2000)
Clitic forms:				
NSubj + ProObj	(0)	85.3% (29)	100% (28)	88.9% (40)
ProObj + NSubj	100% (63)	14.7% (5)	(0)	11.1% (5)
Subtotal	100% (63)	100% (34)	100% (28)	100% (45)
Non-clitic forms:				
NSubj + ProObj	15.7% (43)	75.1% (163)	87.6% (113)	91.0% (181)
ProObj + NSubj	84.3% (231)	24.9% (54)	12.4% (16)	9.0% (18)
Subtotal	100% (274)	100% (217)	100% (129)	100% (199)
Total	100% (337)	100% (251)	100% (157)	100% (244)

Table 6: *Order of Dutch noun subject and personal pronoun object in three centuries based on clitic (reduced) status, with additional 20th century data from Shannon (2000)*

Clearly there is no simple, across-the-board correlation between clitic form and preposing which would hold for all periods examined. In

¹⁷ It cannot be ruled out that some of the written full forms actually concealed reduced clitics. This may mean that a few clandestine clitics were included in the nonclitic count (especially in the later centuries), but their numbers do not appear to be large. Incidentally, Gestel et al. (1992: ch. 4) adopt a nontraditional definition of cliticization, according to which it is a syntactic notion, embracing elements which have been repositioned by a movement rule. In their view reduction is a separate phenomenon.

the 16th century clitic object pronouns such as *-(e)t* ‘it’, *-en* ‘him’, *-se* ‘her/them’ do appear to obey Wackernagel’s law completely: they are all reduced, appear right after and – with a single exception where enclitic *-er* ‘there’ intervenes – are in fact written together with their host, which is the first prong. However, even among the nonclitic forms, almost 85% are preposed and also appear right after the first prong, though none are found written together with it. Nonetheless, according to chi-square, the difference between clitic (reduced) and nonclitic forms is significant ($< .01$). Later, however, the situation is quite different. By the 19th century, weak forms are not written together with their potential host and no longer obligatorily appear right after the first prong. In fact, they are rarely preposed in our data, only (15%), which is somewhat less often than nonclitic forms (25%), though the difference is not significant, according to chi-square. Our 20th century corpus does not contain a single example of a preposed clitic, but this is clearly a fluke, as there were five examples of clitic preposing out of forty-five in the corpus gathered for Shannon (2000). Combining the 20th century data, chi-square again reveals no significant difference in preposing between clitic and nonclitic forms. Below we give a few representative examples.

- (16) a. 16c Doen namse *Turias* in sijn armen, ... [L 7]
 ‘Then Turias took her in his arms, ...’
 b. 16c In Gabaa vertoochde **hem mij** *God* ende vervolde mij
 met wijsheden. [V 18]
 ‘In Gabaa God presented himself to me and filled me with
 wisdoms.’
- (17) a. 19c Zó vond **haar** *den Ekster* tot zijn grote vreugde. [D 25]
 ‘Thus den Ekster found her to his great joy.’
 b. 19c Stom ... zagen *de beide mannen ze* in de duisternis ver-
 dwijnen [B 17]
 ‘Silently ... both men saw them disappear in the dark-
 ness.’
- (18) a. 20c “Er staan **je** *grootse belevnissen* te wachten,” onthulde
 Ongering, ... [Hm 349]
 ““There are great experiences awaiting you.” Ongering
 revealed, ...’

- b. 20c “Ja, die gaf oom Robert **me**,” legde ze uit. [He 48]
 “‘Yes, those Uncle Robert gave me,’ she explained.’”

We conclude that while clitics were obligatorily preposed (and presumably hosted by the first prong, or C) up through at least the 16th century in Dutch, by the 19th century this situation was dramatically changed. In the 19th and 20th centuries, both clitic and nonclitic forms were much more frequently postposed than preposed. Interestingly, even though clitics originally occurred only preposed, they have not shown greater resistance to postposing than nonclitics. In fact, in some cases clitics may have even been less resistant to postposing than nonclitics (cf. 17b; §4.3).

2.2.2. Linear order and subject noun properties

Next we consider correlations between subject properties and the order of noun subject and pronoun object. The properties considered here are: subject type (common vs. proper), definiteness, animacy, and semantic role. Again only examples which contain personal and reflexive pronouns are considered.

2.2.2.1. Subject type. First we look at the correlation between subject noun postposing and the type of the noun subject, common vs. proper. Table 7 below gives the breakdown.

	16 th c.	19 th c.	20 th c.
Common Noun			
NSubj + ProObj	11.8% (34)	82.5% (156)	82.5% (174)
ProObj + NSubj	88.2% (255)	17.5% (96)	17.5% (37)
Subtotal	100% (289)	100% (252)	100% (211)
Proper Noun			
NSubj + ProObj	25.0% (27)	94.2% (113)	98.4% (61)
ProObj + NSubj	75.0% (81)	5.8% (7)	1.6% (1)
Subtotal	100% (108)	100% (120)	100% (62)
Total	100% (397)	100% (372)	100% (273)

Table 7: *Order of Dutch noun subject and pronoun object in three centuries as a function of subject noun type (only personal and reflexive pronoun objects)*

We see here that there is a decidedly weaker proclivity for proper nouns to occur after a pronominal object – as opposed to common nouns, which in all three centuries display a much stronger tendency to postpose. Chi-square indicates significant differences: for all three centuries $p < .01$, at least. Presumably this difference is due to the fact that proper nouns are quite short (usually one word long), human, and contextually given (in fact, often a major protagonist in the narrative), whereas common nouns may not be. Furthermore, although with proper noun subjects postposing was quite common in the 16th century, by the 19th few cases of such postposing were left, and in modern Dutch proper noun subjects are almost never postposed. In Shannon (2000) not a single example out of 164 had a postposed proper noun subject, and only one was found in the current study. Thus, out of a total of 226 cases in the two studies combined, only a single example of a postposed proper noun subject was found – clearly this is an all but extinct pattern (0.4%). The one case we found (21b) is included with a few other examples below; note that the subject is quite long (compound) and actually contains a common noun as the last conjunct.

- (19) a. 16c Ende die coninc van vrancrije en mach niet eten voort
hem zelyadona heet. [S 100]
 ‘And the King of France may not eat before Zelyadona allows him [it].’
 b. 16c Als *Turias* **hem dat** hoorde seggen, was hi seer verwonderd ... [L 71]
 ‘When Turias heard him say that he was very amazed ...’
- (20) a. 19c “Waar gaat ge?” riepen **haar Witte Manse en de andere vrouwen** achterna. [B 35]
 “‘Where are you going?’ Witte Manse and the other woman called after her.’
 b. 19c “Hoe maakt *Louis* ‘**t**, Margo?” vraagt Victorine. [Hm 133]
 “‘How is Louis doing [it], Margo?’ Victorine asks.’
- (21) a. 20c De volgende dag had *d’India* **me** een gedicht gegeven, ... [N 60]
 ‘The next day d’India had given me a poem, ...’

- b. 20c Onder de Farizeeërs en andere lafaards van Joden bevinden **zich** *Achiel de postbode, Rik de bakker en verschillende jongens uit zijn klas*. [C 22]
 ‘Among the Pharisees and other Jewish cowards were Achiel the mailman, Rik the baker, and various boys from his class.’

2.2.2.2. Subject definiteness. Next we examine linear order as a function of the (in)definiteness of the subject noun. Those data are given in table 8.

	16 th c.	19 th c.	20 th c.
Definite Subject			
NSubj + ProObj	15.5% (58)	79.0% (249)	91.5% (204)
ProObj + NSubj	84.5% (316)	21.0% (66)	8.5% (19)
Subtotal	100% (374)	100% (315)	100% (223)
Indefinite Subject			
NSubj + ProObj	13.0% (3)	35.0% (20)	62.0% (31)
ProObj + NSubj	87.0% (20)	65.0% (37)	38.0% (19)
Subtotal	100% (23)	100% (57)	100% (50)
Total	100% (397)	100% (372)	100% (273)

Table 8. *Order of Dutch noun subject and pronoun object in three centuries as a function of subject definiteness (only personal and reflexive pronoun objects)*

In keeping with a widely cited tendency for subjects to refer to topical entities, definite subjects are at all times much more frequent than indefinites, but the percentage of indefinite subjects increases over time here, from 6% (23/397) in the 16th century to 15% (57/372) in the 19th century to 18% (50/273) in the 20th. The reason for this increase is not clear, but may well have more to do with differing narrative styles and conventions than with any purely structural factors.

More importantly for our present concerns, while in the 16th century there is hardly any difference in noun subject postposing as a function of subject definiteness (chi-square shows no significance; both types of subject very frequently display the predominant postposed order), indefinite subjects later are clearly more prone to postposing than definites (confirmed by chi-square: for both centuries $p < .0001$), and this

tendency is even more pronounced in the 20th century than in the 19th. In the latter century subject noun postposing is in fact almost reversed in frequency for indefinite subject nouns (65%) as opposed to definites (21%), whereas in the 20th century although subject postposing is under 10% with definite subjects, it is still comparatively alive with indefinites, where it is found in almost 40% of the relevant cases. If we adjust for the effect of proper noun subjects by removing them from our count, the difference is absent in the 16th century data, but still present for the 19th and 20th centuries, although it is somewhat less pronounced for the 19th, where postposing of definite subjects rises some when only common noun subjects are considered. The relevant adjusted figures for postposed definite common noun subjects are: 16c 88% (235/266), 19c 35% (59/195), 20c 11% (18/162). Note finally that when the construction clearly marks the subject as rhematic, such as with “presentative *er*” sentences (cf. Kirsner 1979, Klooster 2000), the subject is (almost) always indefinite and postposed, thereby confirming that clearly rhematic subjects favor postposing after a pronoun object (cf. 18a above, 29b below). In fact, presentative subjects typically occur very late in the middle field.

2.2.2.3. Subject animacy. In addition, the animacy of the subject was examined; cf. table 9.

	16 th century	19 th century	20 th century
+anim subject			
NSubj + ProObj	15.7% (55)	88.5% (201)	95.5% (168)
ProObj + NSubj	84.3% (296)	11.5% (26)	4.5% (8)
Subtotal	100% (351)	100% (227)	100% (176)
-anim subject			
NSubj + ProObj	13.0% (6)	46.9% (68)	69.1% (67)
ProObj + NSubj	87.0% (40)	53.1% (77)	30.9% (30)
Subtotal	100% (46)	100% (145)	100% (97)
Total	100% (397)	100% (372)	100% (273)

Table 9: *Order of Dutch noun subject and pronoun object for three centuries as a function of subject animacy (only personal and reflexive pronoun objects)*

First of all, in keeping with an often noted tendency, subjects strongly tend to be animate: 16c 88% (351/397), 19c 61% (227/382), 20c

64% (176/273). Furthermore, for all periods considered, animate subject nouns are less often postposed than inanimates. While the difference in our 16th century data is minimal (3%) – and not significant, according to chi-square – for the 19th century it is huge (over 40%!), and even in the 20th century it is still considerable (ca. 25%) – in modern Dutch only inanimate subject nouns postpose often (31%), while animates almost never do (> 5%). The differences in the latter two centuries are statistically significant, according to chi-square (in both cases *p* is well below .0001). When we compare these data with those reported above on pronouns (§2.2.1.3), we find that both subject nouns and object pronouns manifest the same basic tendency for animates to appear earlier in the middle field, inanimates later.

Moreover, when both noun subject and pronoun object are human, the noun subject is even less frequently preposed: 16c 79% (204/257), 19c 11% (12/108), 20c 0% (0/76!). However, even in the 20th century subject noun postposing is not impossible here: there were a few such cases of subject postposing (3/128) in Shannon (2000). Nevertheless it is clear that subject noun postposing is almost never found when both the subject and object are human: the combined corpora display postposing under these circumstances only about 1.5% of the time (3/204). Incidentally, the three cases of subject postposing from Shannon (2000) all involved nonagentive subjects (see the next section on the semantic role of the subject). Cf. the following example.

- (22) 20c En toen kwam **hem** ineens *die meid* weer voor de geest, ...
 [Hs 331]
 ‘And then suddenly that girl appeared before his mind, ...’

2.2.2.4. Subject semantic role. Finally we consider the possible correlation of linear order with the semantic role of the subject – agentive vs. nonagentive.¹⁸ The results are given below.

¹⁸ In general such determinations were rather easy to make, but there were a small number of cases where judgments were not as firm. When in doubt I tended to assign agentive. Moreover, in a few cases the semantic role was left open. In general, however, semantic role assignments were relatively unproblematic.

	16 th c.	19 th c.	20 th c.
+agent subject			
NSubj + ProObj	16.9% (58)	79.5% (240)	91.2% (217)
ProObj + NSubj	83.1% (285)	20.5% (62)	8.8% (21)
Subtotal	100% (343)	100% (302)	100% (238)
–agent subject			
NSubj + ProObj	6.0% (3)	39.7% (27)	48.5% (16)
ProObj + NSubj	94% (47)	60.3% (41)	51.5% (17)
Subtotal	100% (50)	100% (68)	100% (33)
Total	100% (393)	100% (370)	100% (271)

Table 10: *Order of Dutch noun subject and pronoun object for three centuries as a function of subject agentivity (only personal and reflexive pronoun objects)*

We observe first that subjects overwhelmingly tend to be agentive (the prototypical subject; cf. Shannon 1987 and references there). Furthermore, while agentive subjects are almost exclusively animate, nonagentive ones are very often inanimate. Most importantly for our present interests, there appears to be a pervasive correlation between the semantic role of the subject and linear order. Non-agentive subjects clearly tend more toward postposing than agentive subjects, in all three centuries. While agentive subject noun postposing drops very rapidly from 84% to 21% to 9%, with nonagentive subject nouns the decrease is not nearly as severe: 94% > 60% > 52%. According to chi-square the 16th century data just miss the $p < .05$ level, while for the other two centuries the differences are highly significant ($p < .00001$). Clearly, postposing for (human) agentive subjects has been decreasing over time and when the object is also human, it is apparently out in modern Dutch (cf. §2.2.2.3). Neither the corpus for 20th century Dutch assembled for this study nor the one for Shannon (2000) contains a single example of this type. Texts from the 19th century still had this now highly marked word order from time to time, as in (17a above; cf. §3), but nowadays this order is apparently unknown.

Next let us consider briefly the semantics of typical predicates displaying nonagentive subjects. Space precludes the longer discussion that this topic deserves here, but we should at least highlight some of the main types. The subjects of such predicates are predominantly inanimate, and the predicate typically represents a situation outside the (volitional)

control of the subject (or object). With personal pronoun objects, the situation in question is often an emotional (24b, 25b), mental (23a, 24a, 25c), or other experience beyond the object's control (23b; 18a), including the approaching or passing of something/someone, physically or mentally (24c, 25c; 22 above). Typically, in these cases, the object is some sort of human experiencer, the subject an inanimate stimulus. Passives of transitive verbs with an indirect object also fit in here (23c, 25a), as they often show preposing of the object pronoun. Given that the human experiencers in such examples are often datives, this may in fact be the source of some speakers' intuition that object preposing is favored with "indirect objects" (cf. §2.2.1.2).

- (23) a. 16c ... daer om bevallen **u** *der vrouwen seden* ... [V 28]
 '... for that reason women's manners please you ...'
 b. 16c ... so dat **hem** noyt *druppel bloets* wt den live en liep. [L 41]
 '... so that never [a] drop of blood ran (to him) from his body.'
 c. 16c ... so was **hem** *groote ere* ghedaen ende groote weerde. [S 95]
 '... so was done (bestowed) to him great honor and great esteem.'
- (24) a. 19c ... lijkt **me** *de therapeutische of verlossende werking hiervan*, zeer gering. [Hm 59]
 '... its therapeutic or freeing effect seems to me very small.'
 b. 19c ... becroop **haar** opnieuw 'n *weelderig gevoel van onrustig jong leven*. [He 39]
 '... a rampant feeling of restless young life overcame her again.'
 c. 19c Onwillekeurig ontsnapte **haar** *een zucht* bij zijn lof. [D 152]
 'Involuntarily a sigh escaped her [she let out a sigh] at [on hearing] his praise.'

- (25) a. 20c ... dat vader een kunstenaar is, ook al is **hem** *Kaste Vier* toegewezen. [G 21]
 ‘... that father is an artist, even if he has been assigned to Caste Four.’
- b. 20c Even liepen **hem** net als vroeger *de rillingen* over de rug. [Gi 124]
 ‘Just then shivers ran (to him) up his back just as earlier.’
- c. 20c ... dat **ons** wel, vanaf een pleintje, *een verraderlijk schijnsel* naderde, ... [Mo 28]
 ‘... that from a little square a treacherous figure approached us ...’

With reflexive pronouns – particularly in modern Dutch – the subject is again regularly inanimate and the predicate often denotes the location of an entity, or its coming into existence, onto the scene, occurring or unfolding (26a, 27a, 27b, 28a; 21b above). Furthermore, involuntary experiences are still to a certain extent still found with preposed object pronouns (28b; cf. above on personal pronoun objects). Formerly, internal (emotional) reactions (26b) or controlled behavior (16b) with animate subjects were also found with preposed reflexives, but such examples appear to be less common nowadays. Observe that long subjects – e.g. ones modified by a(n) extraposed relative clause – are not uncommon here.

- (26) a. 16c Als **hem** *den dach* verbaerde so was heyndric van lymborch in die were. ... [S 108]
 ‘When day came [lit. ‘revealed itself’] Hendrik van Limburg was busy ...’
- b. 16c Doen verblijde **hem** *die bisscop* boven maten. [S 37]
 ‘Then the bishop was extremely happy.’
- (27) a. 19c Tegen de witte gordijnen teekende **zich** *het donkere silhouette van iemand* af. [P 21]
 ‘Against the white curtains the dark silhouette of someone stood out.’
- b. 19c Daar waar **zich** nu *de hoofddeur* bevond, had het bestaan. [P 104]
 ‘There where the main entrance was located it [his house] had stood.’

- (28) a. 20c Maar er heeft **zich** gisteravond *een incident* voorgedaan
dat ... [Gi 81]
 ‘But there occurred an incident last night that ...’
- b. 20c Plotseling wrong **zich** *een diepe zucht* uit het doodstille
 lichaam omhoog. [Gi 155]
 ‘Suddenly a deep sigh wrestled up from the deathly still
 body.’

Again, in modern Dutch, clauses with such predicates are the ones in which the subject is most often postposed (52%). Their subject is nonagentive, typically an uncontrolled inanimate entity or situation. One other characteristic feature of all the predicates in these groups is that very often the subject is unpredictable, unexpected, unforeseen. That is, often it is the subject that is the rhematic, focused new element in the clause, which is being introduced into the discourse at this point (though it does not have to continue to be the topic of the following discourse, it may). This is presumably the reason why subjects in such cases are often quite lengthy (5+ words: 24a, 24b, 27a, 28a; cf. 30a–c below). Both of these factors are linked to subject postposing. Moreover, with personal pronouns the asymmetry of the semantics (animate object, inanimate subject) makes subject assignment relatively unambiguous (cf. §4.3)

2.2.3. Order and length difference

Finally, we consider word length. This factor is of particular interest because it is often claimed that length plays a role in linearization – to wit, longer elements tend to follow shorter ones. Cf. Behaghel’s “Gesetz der wachsenden Glieder” (1932), as well as the late Simon Dik’s “language-independent preferred order of constituents.” Jack Hawkins’ (1994) principle of Early Immediate Constituents also predicts for languages such as ours that short should precede long, since that order provides clear advantages for on-line parsing. Here are first of all the average word lengths of nominal subjects and pronominal objects, as related to the definiteness of the subject noun.

	16 th c.	19 th c.	20 th c.
Definite Noun	NSubj/ ProObj	NSubj/ ProObj	NSubj/ ProObj
NSubj + ProObj	2.36/1.07	2.47/1.04	2.55/1.00
ProObj + NSubj	2.33/1.00	5.42/1.00	7.42/1.00
Subtotal			
Indefinite Noun	NSubj/ ProObj	NSubj/ ProObj	NSubj/ ProObj
NSubj + ProObj	4.67/1.67	4.80/1.00	2.94/1.00
ProObj + NSubj	2.45/1.00	4.38/1.00	6.74/1.00
Subtotal			
Total	NSubj/ ProObj	NSubj/ ProObj	NSubj/ ProObj
NSubj + ProObj	2.40/1.10	2.65/1.03	2.60/1.00
ProObj + NSubj	2.33/1.00	5.05/1.00	7.08/1.00

Table 11: *Average word length of Dutch noun subject and pronoun object in three centuries as a function of word order and subject noun definiteness (only personal and reflexive pronoun objects)*

Several interesting – though for the most part not surprising – correlations can be noted here. First of all, nominal subjects are on average longer than pronominal objects, which overwhelmingly tend to be just one word long; the difference in average length is always at least by a factor of two. Moreover, indefinite subjects tend to be longer than definites – with the sole exception of the 16th century, where the data are too scant to permit reliable conclusions (the average length of 4.67 with preposed indefinite subject nouns is based on three examples, and there were only 23 cases of indefinite subjects).

More importantly in our context, postposed nominal subjects tend to be longer than preposed ones (again with the exception of the 16th century), and the length difference increases steadily (overall: $0 > 4 > 6$). However, the difference in average length of post- vs. preposed subject nouns is much more profound with definite subjects than with indefinites. With definite subject nouns, the length differences of postposed over preposed subjects were: 16c -0.03, 19c 2.95. 20c 4.87; with indefinite subjects they were: 16c -2.22, 19c -0.42. 20c 3.8. In fact, with indefinites, it is not until the 20th century that postposed subject nouns are on average longer than preposed ones, whereas with definite subject nouns the length difference between post- and preposed subjects is clear

already in the 19th century and in each century it is always greater than the corresponding length difference with indefinites. What this seems to indicate is that over time subjects need inter alia to be ever longer in order to be postposed, especially if they are definite. This is particularly clear in the 20th century corpus, where both definite and indefinite postposed subjects are on average more than twice as long as preposed subjects of the same type.

We can also approach this issue by considering the frequency of pre- vs. postposing depending on the word length difference between subject and object. Again, Hawkins' (1994) principle of Early Immediate Constituents predicts for languages such as ours that short elements like pronoun objects should precede long ones like noun subjects, and furthermore that this effect should become more pronounced the greater the length difference. Table 12 gives the breakdown for this comparison. We only consider those examples where the subject was longer than or equal to the object in word length; there were only a handful of instances where that was not the case.

difference →	0	1	2	3	4	5	6	7	8	9	10+	Total
16th century												
NSubj + ProObj	16	25	5	6	1	4	0	0	0	0	1	58
ProObj + NSubj	83	184	38	8	13	2	2	1	2	0	3	336
Total	99	209	43	14	14	6	2	1	2	0	3	394
19th century												
NSubj + ProObj	89	97	32	12	19	1	6	2	1	1	7	267
ProObj + NSubj	5	36	14	11	8	6	6	1	3	0	13	103
Total	94	133	46	23	27	7	12	3	4	1	20	370
20th century												
NSubj + ProObj	48	120	27	10	11	7	3	5	2	0	2	235
ProObj + NSubj	0	12	4	0	1	0	3	5	3	3	7	38
Total	48	132	31	10	12	7	6	10	5	3	9	273

Table 12. *Order of Dutch noun subject and pronoun object in three centuries as a function of word length difference (noun subject length \geq pronoun object length; only personal and reflexive pronoun objects)*

Again we see a correlation between word length and linearization, in fact the same tendency for longer subjects to appear postposed. Except for the 16th century – where the norm was postposing of the nominal subject, and length does not seem to correlate strongly with subject

postposing – the clear trend is that the longer the subject the greater the frequency of postposing. When we group these data together with an interval of four, the frequency of subject postposing increases for the 19th and 20th centuries monotonically and quite drastically as a function of increased difference in length between the subject and object: 16c: 0–3 86% (313/365), 4–7 78% (18/23), 8+ 83% (5/6); 19c: 0–3 22% (66/296), 4–7 43% (21/49), 8+ 64% (16/25); 20c: 0–3 7% (16/221), 4–7 26% (9/35), 8+ 76% (13/17). Clearly, then, the frequency of subject postposing correlates with the length difference between subject and object, as one would expect from the functional principles cited above, especially from Hawkins’ EIC predictions. However, it appears that when length is a factor in subject noun postposing, the length difference must be ever greater when postposing occurs, so that in the 20th century, only really long subject nouns frequently get postposed, especially if they are definite and/or human. Cf. the following.

- (29) a. 20c Uit het groepje mannen maakte **zich** *dezelfde jongen* los
die al eerder had spoken. [Gi 142]
 ‘From the little group of men the same boy broke away
 who had spoken earlier.’
- b. 20c Natuurlijk wilden ze geloven dat er **zich** in dat lompe,
 bonkige omhulsel *een koninklijke, onzichtbare, onsterfe-
 lijke substantie* ophield, *die geen substantie was, iets dat*,
 ... [N 70]
 ‘Naturally they wanted to believe that in that bulky,
 scrawny shell a royal, invisible, immortal substance
 dwelled which was not a substance, something that ...’
- c. 20c ... staarde **haar** uit de bloemen opnieuw *die twee zwarte
 gaten aan, die maar niet gevuld schenen te willen raken*.
 [He 81]
 ‘...the two black holes, which didn’t seem to want to be
 filled, stared at her again from the flowers.’

3. Summary

The results reported above indicate substantial syntactic changes in Dutch over the past 500 years. It is quite certain that there has been a major shift over time in the favored linearization of noun subjects and

pronominal objects in the Dutch middle field. Whereas at least through the 16th century pronoun object preposing was clearly the most frequent order, in the 19th century the opposite was true, and by the 20th century the older preferred order had all but disappeared. We also found that a number of properties of subject noun and object pronoun correlated with the pre- vs. postposing, although in the 16th century object pronoun preposing was so much the rule that the effects of these factors were generally not noticeable. In general one can say that in modern Dutch only personal or reflexive pronoun objects prepose. Moreover, pronoun object preposing is mainly found when factors which favor it are present: in the (personal) pronoun object, human, perhaps experiencer (dative); in the subject noun, (long) indefinite, common noun, inanimate and non-agentive. Any of the opposite properties of subject and/or object appear to disfavor object preposing, and a greater number of the factors in one direction seems to have a cumulative effect on object preposing. Finally, it should be noted that these properties often tend to cluster.

On the whole, our results for 20th century Dutch are quite comparable to those reported in Shannon (1997) and (2000). In particular, all three studies agree on the whole both on the relative infrequency of object pronoun preposing in modern Dutch as well as on the factors which (dis)favor it. Moreover, in comparing Dutch with German Shannon (2000) points out that while in modern German object pronoun preposing is still the statistical rule, in general the same factors at work in Dutch (dis)favoring object preposing are operative in German as well. In addition, the 20th century findings in these studies compare quite well with those in Nieuwborg (1968, 1973), at least in terms of the factors which correlate with the competing orders. However, in Nieuwborg's studies the reported frequencies of object pronoun preposing are somewhat higher. In all likelihood these discrepancies reflect differences in text types, date, and provenance.

Finally, the evidence presented here gives strong empirical validation to Van der Horst's claim that in late 19th century Dutch pronoun object preposing was still quite possible in certain types of sentences but that it had fallen into disuse in such contexts 100 years later. Furthermore, we are now in a position to say what types of sentences have become anomalous with pronoun preposing. It is surely no accident that Van der Horst chose precisely the (kinds of) examples he did; we repeat them here for reference.

- (30) a. Als **hem** *de rector* van school stuurt.
 ‘If the rector sends him from school.’
 b. Hoe maken **het** *je zoontjes*?
 ‘How are your sons doing?’, lit. ‘How do it your sons’

Both example sentences have features which nowadays almost certainly preclude pronoun object preposing. For instance, each has a short, definite, human subject noun. At least the first subject can be classified as agentive, and while the argument status with the verb in (30b) is less clear, it does not appear to have a blatantly nonagentive subject. Note that these subject properties were found to be ones that heavily disfavored pronoun object preposing, more so in the 20th than in the 19th century. Furthermore, the object in the first example is human. Recall that in both this study and Shannon (2000) not a single 20th century example of pronoun object preposing was found with a (definite) human agentive subject and a human object pronoun, although some were in fact found in our 19th century material.

These are apparently the type of examples that Van der Horst had in mind when he made his claim. As he correctly observes, we do find in the 19th century examples of pronoun object preposing in such contexts (cf. 31a–c; cf. also 17a, with a proper noun subject!), but rarely, if at all, in the 20th century. Preposing appears to have been rather common after quotes in our 19th century data (cf. 31b). Observe that (31a) is quite parallel to Van der Horst’s first example (30a), as is (31c) to his second (30b). However, even in the 19th century object preposing certainly was not obligatory, nor even highly frequent, with such examples; we also find similar instances where the pronoun object is postposed (31d, e; 20b above). Of course, this is not surprising, given that already in the 19th century there are many cases of object postposing in contexts which favor object preposing, e.g. with inanimate, nonagentive – but definite! – subjects (31f–g).

- (31) a. 19c... zodat **hem** *de koetsier* met luider stem moest roepen
 ... [D 204]
 ‘... so that the coachman had to call him with a louder
 voice ...’

- b. 19c “Zou je niet ’n uurtje gaan slapen vanmiddag?” vroeg
haar *Bronkhorst* ... [D 129]
 “Shouldn’t you sleep for an hour or so this afternoon?”
 Bronkhorst asked her ...’
- c. 19c Hoe maakt ’t *het kind*? [D 21]
 ‘How is the child doing [it]?’
- d. 19c Hoe *de kinderen* ’t maken? [Hm 13]
 ‘How the children are doing [it]?’
- e. 19c “Een zeer geréussseerde,” had *Bronkhorst* **haar** genoemd.
 [D 133]
 “A very successful one,” Bronkhorst had called her.’
- f. 19c Of ‘t *leven van kelnerin* **haar** beviel? [Hm 117]
 ‘Whether the life of a waitress pleased her?’
- g. 19c ...toen *het denkbeeld van haar kind* **haar** eensklaps voor
 de geest kwam. [B 102]
 ‘... when the image of her child suddenly appeared before
 her mind.’

We interpret Van der Horst’s claim more specifically as follows. While in the late 19th century it was still possible to find object pronoun preposing in clauses containing a (short) definite, human, agent-like subject (and possibly also a human object), in the late 20th (or now early 21st) century such clauses practically never show object preposing (at least in the standard northern variety). Even though his claim was made in a somewhat offhand fashion, it turns out that it was right on the money, as the data in this study have empirically demonstrated. Pronoun object preposing in modern Dutch requires considerable priming from the factors which favor it, as outlined earlier, even more so than was the case in the late 19th century.

4. The difficult question: What caused this shift?

In the preceding we have presented considerable empirical evidence that there has been a continuing shift in word order in Dutch over the past 500 years or more. While as late as the 16th century (and presumably much earlier) Dutch (especially personal and reflexive) pronoun objects consistently preceded nominal subjects, by the 19th and 20th centuries the order had been reversed and such objects now regularly follow noun

subjects in the majority of cases. Shannon (2000) observes the same trend for German, albeit at a much slower pace, so that in modern German object pronoun preposing is still the statistical norm. The next logical question is how and why this shift came about. Given that pronoun object preposing presumably is in some sense “natural” (cf. earlier references to Wackernagel, Behaghel, Dik, and especially Hawkins), it is all the more curious that this order would be abandoned. In the following sections we offer a functional account of why and how this change occurred.

4.1. Towards a functional explanation: Typological drift

Following Shannon (1997, 2000), we maintain that the observed change in medial linearization in Dutch can be fruitfully viewed in the larger context of a more general syntactic drift (cf. Sapir 1921). Burridge (1993) has claimed that in Dutch (and West Germanic in general) there has been a long-term shift from pragmatically determined word order – where linearization is influenced largely by contextual factors – to grammatically determined word order – in which ordering is driven by grammatical relations like subject and object (Thompson 1978). Overall she discerns a change from topic prominence to subject prominence as part of the general drift: older stages of Dutch are claimed to be more topic-prominent than their later descendants. As proof of this she cites ostensible topic-prominent constructions, such as the frequent use of left dislocation (also common in our 16th century corpus), so-called double subjects or floating topics (not found in our corpus), accusative-marked subjects (again frequent in our data), and topic-controlled gapping of grammatically nonparallel constituents (also found in our corpus, though not often).

According to Burridge’s scenario, the stabilization of the position of the finite verb, in particular verb second in main clauses (which itself may have been brought about by Wackernagel’s Law), brought about the need to fill the clause-initial position. This in turn had a number of other consequences, among them the rise of obligatory subject pronouns and “dummy” subjects when the syntactic subject is displaced (extraposition, presentatives) or in impersonal (i.e. subjectless) constructions, and the loss of embracing (double) negation. Eventually a shift from topic prominence to subject prominence was the result; specific symptoms of

this shift included the tightening up of the sentence frame and reduction of exbraciation, and the loss of the above-mentioned topic-prominent constructions.

Despite some uncertainty about the status of some of these constructions, we can accept Burridge's claim of a general shift in Dutch over time from topic to subject prominence.¹⁹ Adopting this perspective allows us to make sense of the shift in pronoun object preposing in Dutch observed earlier in this study by placing it in the overall context of this drift. Viewed against this background, the observed switch from pronoun object preposing to postposing is yet another example of this same general typological shift from more pragmatically to more grammatically determined word order. As a more topic-prominent language, Middle Dutch word order was still to a certain extent driven by contextual factors: short, atonic, contextually presupposed medial elements like pronoun objects were typically ordered before longer, less presupposed elements like noun subjects (perhaps originally in part due to rhythmical reasons – Wackernagel's Law). With the shift to subject prominence, linearization of medial arguments became increasingly dependent on the grammatical distinction between subject and object. Accordingly, the subject became more and more restricted in its placement to a position directly after (alternately, before) the first prong, and hence before medial pronoun objects. In the modern language, it is now only when the nongrammatical factors brought out earlier are strong enough to countermand the prominence of the subject that we find a pronoun object ordered before it in Dutch.

4.2. Drift and “invisible hand” explanations

The scenario we have just offered to account for the shift away from object pronoun preposing in Dutch appeals to the notion of drift: the shift

¹⁹ Note that the account we propose for the shift away from pronoun object preposing in Dutch does not hinge crucially on accepting the specific constructions Burridge gives as examples of the more topic-prominent nature of Middle Dutch. Incidentally, based on other data, Abraham (1982) independently came to a similar conclusion that modern Dutch is more a “subject prominent” language, whereas modern German is a more “topic prominent” language.

from pragmatically determined to grammatically determined word order is in some way responsible for the downfall of object preposing. Such a proposal seems to fit the fashionable term “invisible hand explanation” (Keller 1994): there seems to have been an invisible force, call it drift, pushing in the direction indicated over time. As it stands, however, this account unfortunately remains rather nebulous. Essentially it observes certain general tendencies in the language and says the phenomenon in question can be fit into the same overall direction of change, but it does not show how or why this could have taken place. As Keller observes (67–68) the term invisible hand “can mislead those unfamiliar with the term, seemingly referring to something mysterious and obscure.” But in fact this is not how an invisible-hand theory is supposed to work. Keller goes on to explain: “However, the opposite is true. An invisible-hand theory attempts to explain structures and reveal processes, namely those structures which are produced by human beings who do not intend or even notice them, as if they were ‘led by an invisible hand’.”

A key point in an invisible-hand explanation – here of language change – is that the invisible forces driving the phenomenon need to be explicated and a scenario offered which could have brought about the observed result, without speakers realizing it. For surely speakers of Dutch were not conscious “conspirators” in the drift toward subject prominence, in our case reordering medial subjects and objects. If speakers did not consciously follow this drift, what could have motivated them to gradually reduce the frequency of pronoun object preposing over the past five centuries? While the overall direction of the “drift” seems to be from topic to subject prominence, what might have been the specific factor(s) influencing this change directly? Or to stay with the metaphor: how can we make the invisible hand visible?

4.3. The invisible hand: Loss of inflectional marking and resultant ambiguity avoidance

Following Shannon (1997, 2000), I submit that in large part the shift away from medial pronoun object preposing has been driven by the loss of fairly consistent morphological marking of subject vs. object.²⁰ For

²⁰ In addition to the loss of clear inflectional cues to subject and object, there may well have been other related factors involved in the shift away from object

while Middle Dutch still preserved rather good inflectional indications of what was subject and object, later stages of the language do not. In particular, nominal case has been lost in Modern Dutch, case distinctions in the pronouns have been greatly reduced, and changes in the verbal system led to much reduced morphological cues about subject from verbal agreement. With the loss of more robust non-linear morphological cues of subject and object, the chance for grammatical ambiguity as to what was subject and object increased considerably (cf. Shannon 2000: 180–184 for details). Due to possible confusion as to what was subject, speakers often avoided the potential misinterpretation that preposed pronoun objects could give rise to.

A frequently invoked communicative strategy was to postpone the object in those cases where ambiguity could arise. Note that the placement of a noun subject before a pronoun object is not normally ambiguous with respect to grammatical relations, even in the absence of inflectional cues, but the opposite order is potentially ambiguous. If a morphologically ambiguous medial noun appears before a morphologically ambiguous pronoun, the noun will normally be interpreted as the subject, since nominal objects are (almost?) never placed before pronominal subjects. However, a morphologically ambiguous medial pronoun object appearing before a morphologically ambiguous noun could be mistaken for the subject, because subject pronouns are very frequent and medial subject pronouns are typically placed in front of an object noun. Eventually even unambiguous instances came to follow the same pattern as the ambiguous ones as object pronoun postposing became the statistical norm. In this way, the order “NP subject followed by pronominal object” has come to be almost obligatory in modern Dutch.

This scenario opens, so to speak, the invisible hand of this drift because it shows how and why Dutch speakers more and more came to rely on word order to differentiate subjects from objects without consciously realizing it. The invisible hand behind this drift was the desire to avoid the syntactic ambiguity regarding subject and object which resulted from the erosion of inflectional cues. Speakers did not know that they were

pronoun preposing. For example, as we noted earlier in several places, the frequency of medial objects cooccurring with medial subjects has gone down considerably, especially with datives. Cf. Shannon (2000) for more on this and other factors.

involved in a language drift, rather they were simply acting as good, cooperative interlocutors and tried to avoid misunderstanding of their communicative intent. Dutch speakers avoided syntactic ambiguity over the course of time, which led to the decline of the ambiguous order of pronominal object before nominal subject. In all probability the increasing frequency of preposed subject nouns in turn acted to further strengthen this effect. The change did not take place over night, but rather was the result of large numbers of such individual decisions to postpose a pronoun object when ambiguity could arise, just as the invisible-hand theory would have it.

Ambiguity and its avoidance can also help us to understand a number of the results reported earlier. It is no doubt not fortuitous, for instance, that those pronoun types which rarely or never distinguish case even in the 16th century – e.g. demonstratives like *dat* ‘that’ and *dit* ‘this’ – were not as often preposed. Interestingly, in our 19th century data ambiguous clitic forms like *se* ‘she/her’ – which in Middle Dutch were largely unambiguous and always preposed – were almost never preposed; instead, it was the unambiguous full forms like *haar* ‘her’ which occur preposed (cf. 17a, b). Finally (2.2.2.4), when the semantics of the predicate results in an asymmetry between subject (typically nonagentive inanimate) and object (typically human experiencer) and there is hence little or no possible ambiguity as to what is subject, object pronoun preposing is quite frequent in modern Dutch, even in the absence of morphological cues to subjecthood. But where grammatical ambiguity is most difficult to resolve without sole resort to contextual knowledge, e.g. with human agent subjects and human patient objects, object preposing is always avoided. This is certainly not to deny that in modern Dutch other factors such as length, indefiniteness, etc. may also be relevant, but it seems undeniable that ambiguity has been a major force in driving the linearization of medial noun subjects and pronoun objects.

Of course this is not the first time that avoidance of the ambiguity resulting from the loss of inflectional morphology has been appealed to in explaining a shift in word order, though mine is the first explicit attempt to explain the specific shift in question, to my knowledge. This type of account goes back at least to the pioneering work on drift of Eduard Sapir (1921) and has been explicitly proposed by others such as Hawkins (1986) as well. No doubt the most well-known case of this type of appeal is Theo Vennemann’s (1974) typological explanation for the

putative shift in Germanic from SOV to TVX/SVO. While Vennemann's account has been criticized because that shift would not have actually avoided ambiguity, the criticism does not hold for the gradual drift from pronoun object preposing to postposing outlined here, since the change in question does alleviate the ambiguity problem, as just explained.

But if the scenario we have laid out here for the demise of object pronoun preposing is correct, how is it that one finds preservation of apparent nominal case in late 19th century Dutch, as we observed at the outset (7a–c above)? Surely for increasing syntactic ambiguity to change the order of medial subjects and objects over time, as we have claimed, the noun case system must have been considerably eroded by then. In fact, we believe that nominal case was largely lost much earlier, and the apparent remnants as late as the 19th – and even into the early 20th – century were only archaic forms of the written language with little or no real existence in the speech of the time.²¹ This assertion is supported by Van der Wal (1992: 241ff., 290ff.), among others, who claims that in the 19th and even 18th centuries noun case was only preserved as an archaism of the written language under the influence of Latin grammarians. Weijnen (n.d.: 43) even claims that in the 17th century the case system was considerably eroded. Therefore, it is most likely that the nominal inflections displayed in the late 19th century were little more than a conservative archaism, perhaps as anachronistic as the verbal subjunctive form we also found there (cf. 8).

5. Conclusion

In the present study, we have provided detailed *prima facie* empirical evidence in support of Van der Horst's (1995) claim that in the last century Dutch has seen a decline in the ordering of pronominal objects before nominal subjects and were able to define quite exactly what types of examples have fallen into disuse. Moreover, it was shown that this is part of a shift which has a much longer history dating back to the Middle Ages. We then argued that this change fits into the larger typological "drift" from pragmatically determined word order to grammatically

²¹ Cf. *inter alia* Sapir (1921) for similar comments on the vestiges of case in modern English, a remnant of former times still maintained and defended by language purists.

determined word order discerned by Burridge. Finally, it was claimed that the “invisible hand” driving this shift was the desire to avoid syntactic ambiguity resulting from the loss of clear morphosyntactic distinctions between subject and object in Dutch.

One empirical strength of our historical explanation is that it also makes predictions as to what we should expect to find in related languages that have not been considered here. Where morphological cues to subject and object are better preserved, we should expect to find pronoun object preposing still maintained, while in languages where such distinctions have been largely lost, we should find a shift to pronoun object postposing. Shannon (2000, in prep.) documents the correctness of this prediction for German: since German has not lost nearly as much inflectional morphology as Dutch has, it still preserves more pragmatically determined word order. Preliminary analysis of Afrikaans (Shannon 2003a), Yiddish (2003b), and Low German (Shannon & Dewey in prep.) indicates that the predictions hold true for those languages as well. If so, then this further work will offer persuasive corroboration of our functional explanation for the decline of pronoun object preposing in Dutch, and indeed West Germanic in general.

Sources

16th Century

- Ge = Geeraedts, Loek. 1986. *Het volksboek van Ulenspieghel*. Original ca. 1535. Amsterdam: Pelckmans.
- L = Lecoutere, C. & W. L. de Vreese (eds.). 1904. *Een schoon historie van Turias ende Floreta*. Naar den Antwerpschen druk van de weduwe van Jacob van Liesveldt uit het jaar 1554. [*Nederlandsche volksboeken*, 8.] Leiden: E. J. Brill.
- S = Schellart, Franciscus J. 1952. *Volksboek van Margarieta van Lymborch (1516)*. Amsterdam: De Wereldbibliotheek.
- V = Vreese, Willem de & Jan de Vries (eds.). 1941. *Dat dyalogus of twisprake tusschen den wisen coninck Salomon ende Marcolphus*. Naar den Antwerpschen druk van Henrick Eckert van Homberch in het jaar 1501 uitgegeven. [*Nederlandsche volksboeken*, 7.] Leiden: E. J. Brill.

19th Century

- B = Buysse, Cyriel. 1893/1967. *Het recht van de sterkste*. Hasselt: Heideland.
- D = Daum, A. 1889/1964. *Goena-goena*. Amsterdam: Querido.

Hm = Heijermans, Herman. 1893/1904. *Trinette*. Amsterdam: H. J. W. Becht.

P = Prins, Arij. 1885/1925. *Uit het Leven*. Amsterdam: Kampen.

20th Century

C = Claus, Hugo. 1989. *De zwaardvis*. Amsterdam: De Bezige Bij.

Gi = Gijzen, Wim. 1980. *Rissan*. Amsterdam: Meulenhoff.

Hd = Heijden, A. F. Th. van der. 1992. *Weerborstels*. Amsterdam: Querido.

Hs = Hermans, Willem Frederik. 1975. *Onder professoren*. Amsterdam: De Bezige Bij.

Mo = Moor, Margriet de. 1993/2000. *De virtuoos*. Amsterdam: Contact.

Mu = Mulisch, Harry. 1987. *De pupil*. Amsterdam: De Bezige Bij.

N = Nooteboom, Cees. 1991. *Het volgende verhaal*. Amsterdam: De Arbeiderspers.

References

- Abraham, W. (1981). Zur Wortstellungssyntax Deutsch-Niederländisch. In *Logos Semantikos: Studia Linguistica in Honorem Eugenio Coseriu 1921–1981*, 81–101. de Gruyter, Berlin.
- Balk-Smit Duyzentkunst, F. (1979). Chomskys metaforen. *Spektator* 9, 3–13.
- Behaghel, O. (1932). *Deutsche Syntax*. Vol. 4: *Wortstellung*. Winter, Heidelberg.
- Burridge, K. (1993). *Syntactic change in Germanic*. Benjamins, Amsterdam.
- Gestel, F.C. van, J. Nijen Twilhaar, T. Rinkel & F. Weerman (1992). *Oude zinnen: Grammaticale analyse van het Nederlands tussen 1200–1700*. Nijhoff, Leiden.
- Haeseryn, W., et al. (1997²). *Algemene Nederlandse Spraakkunst*. 2 vols. Martinus Nijhoff, Groningen.
- Hawkins, J.A. (1986). *A comparative typology of English and German*. University of Texas Press, Austin.
- Hawkins, J.A. (1994). *A performance theory of order and constituency*. Cambridge University Press, Cambridge.
- Horst, J.M. van der (1995). Taalverandering in de 20ste eeuw. *Ons Erfdeel* 38, 69–74.
- Keller, R. (1994). *On language change: The invisible hand in language*. London: Routledge. English translation of *Sprachwandel* (1990). Francke, Tübingen.
- Kirsner, R.S. (1979). *The problem of presentative sentences in modern Dutch*. North Holland, Amsterdam.
- Klooster, W. (2000). *Er*: Information structure and specificity. In T.F. Shannon & J.P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1997: The Dutch language at the millennium*, 113–132. University Press of America, Lanham, MD.
- Koster, J. (1978). *Locality principles in syntax*. Foris, Dordrecht.
- Lehman, W.P. (1992). *Historical linguistics*. Routledge, London.

- Nieuwborg, E. (1968). *De distributie van het onderwerp en het lijdend voorwerp in het huidige geschreven Nederlands in zijn A.B.-vorm*. Plantyn, Antwerp.
- Nieuwborg, E. (1973). De plaatsing van het substantivisch onderwerp in reflexieve constructies. *Leuvense Bijdragen* 62, 273–283.
- Sapir, E. (1921). *Language. An introduction to the study of speech*. Harcourt, Brace & World, New York. Paperback edition, 1969.
- Shannon, T.F. (1987). On some recent claims of relational grammar. In J. Aske et al. (eds.), *Berkeley Linguistics Society. Proceedings of the thirteenth annual meeting, February 14–16, 1987*, 247–262. Berkeley Linguistics Society, Berkeley.
- Shannon, T.F. (1997). Word order change in Dutch as reflected in the *Ulenpieghel*. *Northwest European Language Evolution* 31/32: 361–88.
- Shannon, T.F. (2000). On the order or (pro)nominal arguments in Dutch and German. In T.F. Shannon & J.P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1997: Dutch Linguistics at the Millennium*, 145–195. University Press of America, Lanham, MD.
- Shannon, T.F. (2003a). Nederlands tussen Duits en ... Afrikaans. Talk presented at the 15th Colloquium Neerlandicum, Groningen, August 28, 2003.
- Shannon, T.F. (2003b). New evidence for drift in Germanic. Talk presented at the session on Language Change held at the Annual Meeting of the Modern Language Association, San Diego, December 29, 2003.
- Shannon, T.F. (in prep.). Word order change in German.
- Shannon, T.F. & T. Dewey (in prep.). Word order change in Low German.
- Thompson, S.A. (1978). Modern English from a typological point of view: Some implications of the function of word order. *Linguistische Berichte* 54, 19–35.
- Vennemann, Th. (1974). Topics, subjects, and word order: From SOV to SVX via TVX. In J.M. Anderson & Ch. Jones (eds.), *Historical linguistics: Proceedings of the First International Conference on Historical Linguistics*, 339–376. North Holland, Amsterdam.
- Verhagen, A. (1986). *Linguistic theory and the function of word order in Dutch*. Foris, Dordrecht/Mouton de Gruyter, Berlin.
- Wal, M.J. van der (1992). *Geschiedenis van het Nederlands*. Spektrum, Utrecht.
- Weijnen, A.A. (not dated). *Zeventiende-eeuwse taal*. 5th printing. Thieme, Zutphen.

A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale

Liesbeth Degand (University of Louvain, Louvain-la-Neuve) &
Henk Pander Maat (Utrecht University)

1. Introduction

It is generally acknowledged that coherence relations and their linguistic markers can express meanings at different “levels” of the discourse. These differences in meaning and use have mainly been accounted for in terms of dichotomies: for instance, external/internal (Halliday & Hasan 1976; Martin 1992), semantic/pragmatic (van Dijk 1979; Moeschler 1989), subject matter/presentational (Mann & Thompson 1988). A very influential account is also Sweetser (1990), who has proposed to distinguish not two, but three domains of use for connectives: the content domain, relevant for (1), the epistemic domain, exemplified in (2), and the speech act domain, illustrated in (3).

- (1) John came back because he loves her.
- (2) John loves her, because he came back.
- (3) What are you doing tonight, because there’s a good movie on.

In previous work (Degand & Pander Maat 1999; Pander Maat & Degand 2001) we have proposed an alternative account of the distribution of causal connectives. Going beyond dichotomous and trichotomous classifications, we have proposed to represent (causal) coherence relations and connectives in a scalar way. This scalar representation reflects the fact that (causal) connectives are not strictly domain specific, but that they nevertheless impose constraints on the contexts in which they can occur, with some contexts being more “natural” than others¹ (see also Degand

¹ Actually, there are some connectives that seem to be rather domain-specific, like Dutch *daardoor* and *doordat* which are highly restricted to the non-volitional domain, but such a domain specificity seems to be the exception rather than the rule.

& Sanders 1999). In addition, a number of causal connectives seems to take an intermediate position between the traditional categories. According to us, this situation is an indication for the need of a scalar perspective on the spectrum reaching from non-volitional causality in the content domain to epistemic and speech act causality.

The scale we have developed is one of *speaker involvement*, on which the inherent expressive power of connectives can be represented. Our hypothesis is that the different causal relations can be ordered along a scale from minimal to maximal *speaker involvement*. Speaker involvement refers to *the degree to which the present speaker is implicitly involved in the construal of the causal relation*. More specifically, speaker involvement increases with the degree to which both the causal relation and the related segments carry assumptions and actions of the present speaker (see below). The different causal relations we distinguish are, in order of increasing Speaker Involvement: causal non-volitional and volitional content relations; causality-based and non-causality based epistemic relations, and causal speech-act relations. Pander Maat & Degand (2001) give a detailed account of this relational ordering on the scale. Here it will suffice to give an overview of the characteristics that come into play when determining the level of Speaker Involvement of a causal relation. Afterwards we will discuss the interaction between a number of Dutch and French causal connectives and the Speaker Involvement scale, and support our claims on the basis of a corpus analysis.

2. Speaker Involvement and causal relations

Four characteristics of coherence relations may enhance the prominence of speaker assumptions in the relation, and thus enhance the level of Speaker Involvement of the relation: The involvement of a protagonist, a lack of isomorphism between the relation and states of affairs in the real world, proximity of the relation to the present speaker and the time of speaking and the implicit realization of the protagonist.

First, causal coherence relations may be characterized in terms of the degree to which they necessarily imply the *subjective involvement* of a conscious participant, which may, but need not be the speaker. For instance: a volitional relation (see 4 below) involves such a participant (an actor), while a non-volitional relation (5) does not. Since conscious

participants may entertain assumptions and may be identified with by the speaker, volitional relations carry a higher degree of SI than non-volitional ones. Epistemic relations (6) are still higher on the scale because they involve beliefs of a concluding protagonist with whom the speaker must share a number of assumptions in order to understand the causal relation² (see examples 2 and 6). In fact the hearer must also share these assumptions to a certain extent (see also Pander Maat & Degand 2001). Finally, maximal SI is reached with speech-act relations in which the protagonist is identical to the Speaker by definition. In a speech-act relation the speech-act is motivated by reference to a situation constituting the reason for it (as in example 3 above).

- (4) He left, because he felt tired.
- (5) The temperature went up, because of the fact that the sun came up.
- (6) The children should be punished, because they have been mean.

Second, a given causal relation may be more or less *isomorphic with a real world causal relation*, that is, a causal relation in which the present speaker has no role. For instance, the volitional relation in (1) represents a real world causal relation. The same causal situation is present in (7) below, but this time it is no longer the primary causal relation expressed. In (7), a reason-consequence relation in the real world is transposed into the mental domain of making inferences, or in other words it is used to base an argumentation by the speaker upon. This is why the epistemic relation in (7) carries a higher degree of SI than the volitional relation in (1), repeated for convenience below.

- (7) John must have come back, since he loves her.
- (1) John came back, because he loves her.

In this kind of epistemic relation, real world causality is not represented, but it continues to impose constraints on the propositional content of the related segments. This is no longer the case for non-causal epistemic relations and speech act relations, which show maximal detachment from real world causality. Two examples of non-causal epistemic relations are

² In most cases, this concluding protagonist *is* the speaker himself.

presented in (8) and (2). In (8) the conclusion does not refer to a probable consequence of the situation referred in the premise, but to an evaluation by the speaker. Since this evaluation is not a real world situation, there can be no real world causality involved in its coming about. Hence, it can only be based on ‘inferential’ causality, operating in the domain of reasoning. In example (2), repeated below for convenience, this movement of detachment from real world causality is even more apparent: in this example, the real world causal relation is ‘turned around’ to yield a pattern of abductive inference. In the empirical study presented below, abductive epistemic relations will be discussed separately from other non-causal epistemic relations, since they will turn out to discriminate between connectives.

- (8) John just told me that he will accept our offer. This is an important message, since we assumed that he would not accept.
 (2) John loves her, because he came back.

Third, a given relation may be placed at *different distances from the present speaker*, and at different distances from the moment of speaking. The closer a given relation is to the present speaker, the more it constitutes a vehicle for the expression of speaker assumptions. After all, the speaker is more likely to accept the general assumptions underlying his own decisions than those underlying other persons’ decisions. Hence, the first-person relation in (10) has a higher degree of SI than the third-person relation in (9). And (11) is even higher in SI because it has present tense.

- (9) He felt tired. He left.
 (10) I felt tired. I left.
 (11) I feel tired. I’m going home.

In more general terms, SI increases when the distance between the present speaker and the protagonist decreases. In non-volitional relations SI is minimal because there is no protagonist to identify with. In volitional relations the situation is as sketched in examples (9-11). In epistemic relations the speaker is the unmarked protagonist and in speech act relations he is the only conceivable protagonist. This decreasing distance between speaker and causal protagonist leads to an increasing SI

from non-volitional to speech-act relations. The same happens with regard to the time distance to the moment of speaking: SI increases when the distance between speaking time and causal situation time decreases. Hence non-volitional and volitional relations are lowest on the SI scale because there are no temporal constraints on the expression of these types of causal relation with respect to the speaking time, i.e. the time of the causal event and the speaking time are independent. Speech act relations are highest on the SI scale because causal situation time (t_2) and speaking time (t_s) obligatory coincide, i.e. a speech-act cannot take place at another time than the speaking time. As for the epistemic relations, t_2 and t_s very often coincide but this is not obligatory the case, i.e. generally (speaker) conclusions are presented to be valid at the time of speaking but this is not necessarily so.

Fourth, the involvement of a conscious participant may vary in *explicitness*. We will illustrate this dimension with reference to the presence of the speaker in an epistemic relation.

- (12) He is Hungarian.
- (13) He is probably Hungarian.
- (14) I think he is Hungarian.

Taken in isolation, (12) expresses a factual statement that is produced outside the domain of the speaker's conceptualizing activity. That is, it presents itself as carrying a minimal degree of SI, because the speaker is only minimally involved in the production of the reported cognition. In contrast, (13) and (14) contain some explicit elements referring to the speaker's perspective. In example (14) the modal adverb *probably* invokes the speaker as the source for the probability judgement. This is even more explicit in (14) where the speaker's perspective *I think* is more or less "objectified". This objectification appears from the fact that it may be referred to by demonstrative pronouns in subsequent utterances. E.g. *that is not true* would refer to the *I think* clause, not to its complement. At this sentence level, (13) has a higher SI than (14), which is still higher than (12).

However, factual statements like (12) can also occur in a discourse context that heavily suggests an interpretation as a speaker conclusion.

For instance, this would be the case if (12) would be preceded by *he can only be Czech or Hungarian and I am sure he is not Czech*. In this case the speaker's inferential activity is part of the relational interpretation, but at the same time it is maximally implicit. Hence, this time it has a maximal degree of SI, higher than (13) and still higher than (14).

In sum, the involvement of the protagonist may be more or less implicitly realized in the discourse. As we proceed on the scale, explicit realization (in volitional relations) gradually changes into possibly implicit realization (in epistemic relations) and to obligatory implicit realization (in speech act relations).

Together the different dimensions mentioned determine the level of SI present in a causal relation³, that is the degree to which the relation is a vehicle for the expression of the speaker's assumptions and activities.

3. Speaker Involvement and causal connectives

We claim that the Speaker Involvement Scale is not only a way to classify coherence relations but that connectives – as prototypical markers of coherence relations – can be analyzed in terms of the scale too. One may expect that a valid classification of coherence relations, which are considered as cognitive entities, has some counterpart in language. After all, there are restrictions on the use of connectives. And there are many correspondences between connectives and the relational features they signal. Even though there is no one-to-one mapping between relations and connectives, the restrictions on the use of connectives imply an organization of the relations they can express; they do not just co-exist as a set of relations on one and the same level.

In previous work (Degand & Pander Maat 1999; Pander Maat & Degand 2001) we have demonstrated that the Speaker Involvement Scale can be used to analyze and classify forward causal connectives in Dutch and French. In particular, we have supported the following claims concerning the relationship between connectives and SI:

³ Of course, some of these dimensions may correlate. For instance, the dimension “identity of the protagonist” correlates with the dimension “realisation of the protagonist”, since third-person-protagonists are more often explicit than first-person-protagonists. However, correlation is not determination.

1. A connective *encodes a certain SI level*, which it contributes to the interpretation of its discourse environment. When this level is too low or too high to be combined with the level allowed for by the discourse environment, the use of the connective is inappropriate.
2. The set of relational environments of a connective can be represented as an area on the SI scale. The distribution of every connective occupies a *contiguous area* on the scale. That is, we do not expect to find a single connective that may express, e.g., volitional and non-causal epistemic relations but not causal epistemic relations.
3. The most frequent causal connectives in a given language should *differ significantly* from each other on the scale. That is, the prime reason for a language to have more than one causal connective is to be able to express several levels of SI.
4. These claims are *cross-linguistically valid* (starting with Dutch and French).

In the remainder of this paper, we will extend this account to three backward causal connectives in Dutch and French: *omdat/parce que* ('because'), *want/car* ('because/for'), *aangezien/puisque* ('since'). By 'backward causal connectives' we refer to connectives which are placed in the antecedent-segment of the relation. This does not always imply that this segment follows the consequent-segment, as has been the case in the examples given so far. In fact, some of the connectives are more or less evenly distributed over the two orders consequent-antecedent and antecedent-consequent (see Table 1 below).

The purposes of our study are twofold. First, we want to check whether the SI scale, which was developed on the basis of forward causal relations/connectives can also be used to describe backward connectives. If this were not the case, the generality of the scale would be doubtful. Second, there is a striking frequency difference between the backward translation "equivalents" in Dutch and French. Figure 1 displays the frequencies of the connectives under investigation in Dutch and French newspaper corpora.⁴

⁴ The corpora we used are electronic transcriptions of two daily newspapers. For Dutch, the 1994 issues of *NRC Handelsblad*, a Dutch national newspaper, provided by the INL (Leiden Institute of Dutch Lexicology) containing about

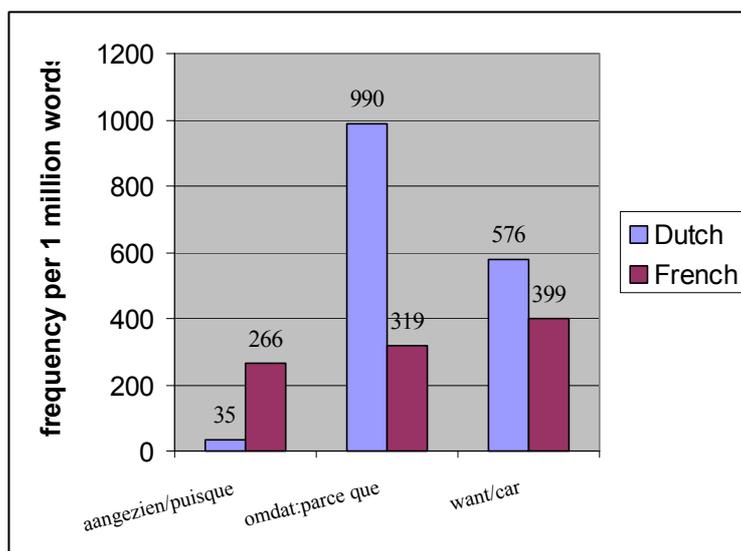


Figure 1: Frequency of Dutch and French causal connectives in newspaper corpora

The most striking frequency divergences occur with *aangezien* and *puisque*, since *puisque* is more than seven times more frequent than *aangezien*. *Omdat* and *parce que* also show highly diverging frequency patterns *omdat* being more than three times more frequent than *parce que* in comparable newspaper corpora. In addition, there is another intriguing difference between the connectives *puisque* and *aangezien*, namely a divergent syntactic pattern with respect to causal segment ordering. In Dutch, *aangezien* occurs nearly as often in antecedent-consequent (*aangezien Q, P*) ordering as in consequent-antecedent ordering (*P aangezien Q*). In French, however, *puisque* occurs nearly exclusively in consequent-antecedent ordering (*P puisque Q*). *Omdat* and *parce que* which also admit the two orderings do not show any divergences here (see Table 1).

27 million words. For French, the 1997 issues of *Le Soir*, a Belgian franco-phone newspaper distributed on CD-ROM containing about 26.8 million words.

	<i>aangezien</i>	<i>puisque</i>	<i>omdat</i>	<i>parce que</i>
P conn Q	27 (54%)	48 (96%)	45 (90%)	45 (90%)
conn Q, P	23 (46%)	2 (4%)	5 (10%)	5 (10%)
	50	50	50	50

Table 1: *Ordering of causal segments*

At the same time, in the traditional linguistic literature (e.g. Groupe λ -1 1975; Iordanskaja 1993; van Belle 1989), these connectives are described in similar terms. For instance, both *puisque* and *aangezien* are described as “utterance markers” rather than “operators”, and the antecedent is presented as already known to the hearer, while the speaker is not supposed to accept the truth of it automatically. This brings us to the question whether there are subtle differences between these backward connectives that could be accounted for in terms of the Speaker Involvement scale, but not on the basis of traditional descriptions.

3.1 A contrastive SI analysis of causal connectives

In our view, analyses of the SI potential inherent to connectives cannot do without systematic corpus analyses. Indeed, while attractive parallels and suggestive similarities between coherence relations and the linguistic devices that express them have been claimed, recent corpus studies do also reveal that existing categorization proposals *cannot* account for the data of connective distribution in a straightforward way (Degand 1998, 2001; Pander Maat 1998; Pander Maat & Sanders 1995). Hence, we carried out corpus analyses of the backward causal connectives in (written) Dutch and French: *omdat/parce que*, *want/car*, *aangezien/puisque*. For each of these connectives, we assembled 50 occurrences from a newspaper corpus.

First, we identified the coherence relation of each fragment by means of a paraphrase test.⁵ The results for Dutch and French are given

⁵ The non-volitional paraphrase was «this has/had the following cause»; the volitional paraphrase was «this action is/was the consequence of the following» and the epistemic paraphrase was «this conclusion follows from the following». For speech-act relations, two kinds of paraphrases were used: «(the use of) this word/phrase is motivated by the following» and «this speech-act is the consequence of/is motivated by the following». Finally, non-causal epistemic rela-

in Tables 2 and 3, respectively. In the two languages, the contiguity hypothesis (the second claim in section 3) is confirmed.

Relation	connective		
	<i>aangezien</i>	<i>want</i>	<i>omdat</i>
non-vol. causal	2 (4%)	1 (2%)	10 (20%)
vol. causal	8 (16%)	7 (14%)	16 (32 %)
causal epistemic	26 (52%)	20 (40%)	20 (40%)
non-causal epistemic	14 (28%)	17 (34%)	4 (8%)
abductive epistemic		1 (2%)	
speech act		4 (8%)	
	50	50	50

Table 2: *Relational interpretations co-occurring with backward causal connectives in a Dutch newspaper corpus*

For Dutch, *aangezien*- and *want*-fragments are clearly higher on the SI-scale in terms of relational interpretations than *omdat*-fragments. The two connectives do indeed differ significantly from *omdat* but not from one another (*aangezien-omdat*: $\chi^2 = 14.34$, $df = 3$ and $p < .01$; *want-omdat*: $\chi^2 = 23.94$, $df = 5$ and $p < .0001$; *aangezien-want*: $\chi^2 = 6.47$, $df = 5$ and $p > .05$; N.S.). This means that the third claim in section 3 is only partially supported in Dutch. However, since *aangezien* is fairly infrequent in Dutch, there is no functional motivation for *aangezien* to be completely distinctive from its backward alternatives. We will discuss this aspect below.

tions were distinguished from causally-based epistemic relations by checking whether substituting a ‘real-world’ causal paraphrase (non-volitional or volitional) results in a coherent sequence (though the meaning differs from the original fragment). If yes, the relation is causally-based, if not it is non-causal.

Relation	connective		
	<i>puisque</i>	<i>car</i>	<i>parce que</i>
non-vol. causal		1 (2%)	11 (22%)
vol. causal	1 (2%)	10 (20%)	23 (46%)
causal epistemic	36 (72%)	27 (54%)	14 (28%)
non-causal epistemic	10 (20%)	9 (18%)	2 (4%)
abductive epistemic	1 (2%)		
speech act	2 (4%)	3 (6%)	
	50	50	50

Table 3: *Relational interpretations co-occurring with backward causal connectives in a French newspaper corpus*

For French, *puisque*-fragments are highest on the SI-scale in terms of relational interpretations, followed by *car*-fragments, which in turn are more subjective than *parce que*-fragments. All connectives differ significantly from one another. (*puisque-car*: $\chi^2 = 12.70$, $df = 6$ and $p < .05$; *puisque-parce que*: $\chi^2 = 49.18$, $df = 6$ and $p < .0001$; *car-parce que*: $\chi^2 = 25.03$, $df = 4$ and $p < .0001$). Hence our third claim is supported for French.⁶

So far, we contrasted the connectives in terms of the relations they occur in. These relational interpretations are the result of an interaction between the connective and the connected discourse segments. But what we are looking for here is the contribution of the connective to this interaction. We have characterized this contribution in terms of ‘a certain SI level that is added to the interpretation of its discourse environment’. Now, how do we determine the SI level encoded by a connective? A first way to proceed would be to postulate that the relational interpretations of the fragments containing a connective also constitute the area on the SI scale covered by the connective, i.e. its specific SI level. However, this line of reasoning only leads to the unsatisfactory conclusion that the SI areas of the different connectives show considerable overlap. For

⁶ Abductive epistemic relations are a specific category of non-causal epistemic relations. Hence, the absence of abductive uses for the connective *car* (Table 3) does not constitute a ‘gap’ in the scale. Furthermore, *car* can be used in such a context. E.g. *Il doit avoir faim, car il a déjà mangé trois sandwiches.* (‘He must be hungry, because he has already eaten three sandwiches’). So, the contiguity hypothesis (claim 2) is also supported for French.

instance, the distinction between causal volitional, causality-based epistemic and non-causality based epistemic relations is not straightforwardly lexicalized for the backward causal connectives, neither in Dutch nor in French.

However, the distribution shows ‘peaks’ at different points. A first clue to the different SI profiles of the connectives may be obtained by inspecting these differences in terms of statistical significance. For Dutch, the difference between *aangezien* and *omdat* was significant, which indicates that *aangezien* may encode a higher SI level than *omdat*. But this approach does not help us in accounting for the differences between *aangezien* and *want*.

We could, of course, make claims based on our own intuitions regarding the meanings of the connectives or based on intuitions expressed in the linguistic literature. However, we prefer a somewhat more cautious approach, which combines distributional data and semantic intuitions.

In order to determine the proper level of SI that a connective contributes to its environment, we propose to analyze the contexts they occur in according to the four SI aspects presented in section 2. That is, (i) the degree of subjective involvement of a conscious participant, (ii) the degree of iconicity of the causal relation, (iii) the distance to speaker and speaking time, (iv) the degree of explicitness of the participants involved in the causal relation. For each occurrence of a given connective, we analyze the related segments for a number of features related to these four aspects. The combination of these distributional data determines the level of SI encoded by the connective. Doing so, we can scale the connectives with respect to each other on the SI scale. From a contrastive point of view, it becomes also possible to compare the respective SI level of “equivalent” connectives in different languages. We will now first proceed with the analysis of the Dutch connectives, then with the French connectives, and finally compare the results in the two languages.

So far, we have established, on the basis of Table 1, that in Dutch the backward connectives *want* and *aangezien* are higher on the SI scale than *omdat*. There is indeed a significant tendency for *omdat* to express causal relations that are isomorphic with real-world causality, i.e. non-volitional, volitional, and causality-based epistemic relations, and which are at a certain distance from speaker and speaking time (non-volitional and volitional relations). These two characteristics place *omdat* lower on

the SI scale than *want* and *aangezien*. The latter connectives can indeed also express relations that are non-iconic with respect to real-world causality (non-causality based epistemic relations), with *want* occurring also in speech-act contexts, i.e. with minimal distance to speaker and speaking time.

Nevertheless, *want* and *aangezien* differ too. According to our intuitions *aangezien* cannot express speech-act relations (see example (16)). Moreover, its acceptability is also doubtful in abductive epistemic relations, especially when the premise precedes the conclusion (see 17))

- (16) a. ?Kom onmiddellijk naar binnen, *aangezien* het regent.
 ?Come inside immediately, [*aangezien*] it is raining.
 b. ?Wat doe je vanavond, *aangezien* er een leuke film draait.
 ?What are you doing tonight, [*aangezien*] there is a good movie on.
- (17) a. ?Hij rookt, *aangezien* hij sigaretten koopt.
 ?He smokes, [*aangezien*] he is buying cigarettes.
 b. **Aangezien* hij sigaretten koopt, rookt hij.
 **[Aangezien]* he is buying cigarettes, he smokes.

Hence, we hypothesize that *want* is higher on the SI scale than *aangezien*. This should be confirmed or disconfirmed on the basis of further distributional data. We will further investigate the difference in SI between *want* en *aangezien* by focussing on two SI dimensions that have been mentioned already in the introduction: the distance between the relation and the present speaker and the realization of the protagonist of the relation. To uncover these SI aspects, we proceed in two steps.

First we determine the identity of the causal protagonist in the first segment (S1). To this end, we first select the fragments which actually have a causal protagonist, thus leaving out the factual segments, and which show some variation with respect to the identity of this protagonist, thus excluding speech-act fragments for which the identity of the causal protagonist is always the speaker. This leaves us with volitional and epistemic relations for which the possible participants include the author, a group including the author and others, a cited speaker, a generic third person, a pronominal third person and a nominal third person. The principal distinction in this domain is the one between first-person participants and third-person participants (we don't find

second person protagonists in our newspaper corpora). The hypothesis is that a connective with a higher SI level more often occurs with first-person participants, since the actions and conclusions of speakers will generally be formulated with a higher SI level than the actions and conclusions of third persons. After all, the speaker is more likely to accept the general assumptions underlying his own decisions than those underlying other persons' decisions.

The second step is to determine how this responsible participant is realized in the consequence segment (S1), i.e. whether he is linguistically expressed or not, and if this is the case whether this protagonist is explicitly (*I think/ he thinks*) or implicitly (e.g. *probably*) present in the segment. In non-volitional relations there is no conscious causal participant, so the question does not arise. In volitional relations, this participant is nearly always explicitly mentioned (with the exception of some passive and impersonal constructions), so the question arises mainly in epistemic relations. For these relations, we first compare the number of factually presented consequence segments (no participant involved) with the number of subjective consequence segments. Within this latter category we distinguish the implicit from the explicit cases.

Of course, this latter distinction is not entirely independent of the distinction between first and third person protagonists, since generally first persons are implicit and third persons explicit. However, since first persons are occasionally explicit and third persons occasionally implicit, it makes sense to formulate a separate hypothesis regarding implicit and explicit realization of the participant responsible for the causal relation. Since we have defined SI as the degree of implicit involvement of the present speaker in the relation, our hypothesis must be that a more subjective connective more often occurs in first segments with implicit participants. When no participant is involved in the segment (factual conclusions in epistemic relations), the SI hypothesis predicts a higher occurrence of high-level SI connectives, since these connectives are capable of introducing subjectivity into the relational interpretation all by themselves.

The results regarding these three hypotheses on the difference between the Dutch forward connectives are presented in Figures 2 and 3, and in Table 4. Since we are most interested here in the contrast between the connectives *want* and *aangezien*, results for these connectives are given separately in addition to the global results.

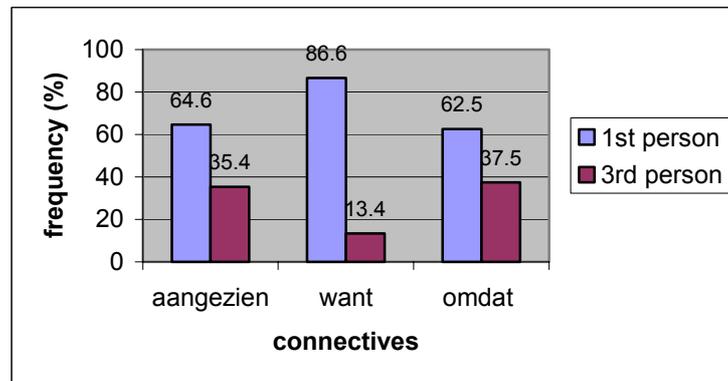


Figure 2: Identity of the causal participant co-occurring with aangezien, want or omdat in volitional and epistemic relations
 $\chi^2 = 7.785$, $df = 2$, $p < .05$ ($p = .020$)
 want-aangezien: $\chi^2 = 6.085$, $df = 2$, $p < .05$ ($p = .014$)

Participant	connective		
	aangezien	want	omdat
Subjective consequence	40 (100%)	37 (97.4%)	24 (100%)
Factual consequence	0	1 (2.6%)	0
	40	38	24

Table 4: Type of consequence segment co-occurring with aangezien, want or omdat in epistemic relations

$\chi^2 = 1.701$, $df = 2$, $p > .05$ ($p = .427$) N.S.
 want-aangezien: $\chi^2 = 1.066$, $df = 1$, $p > .05$ ($p = .302$) N.S.

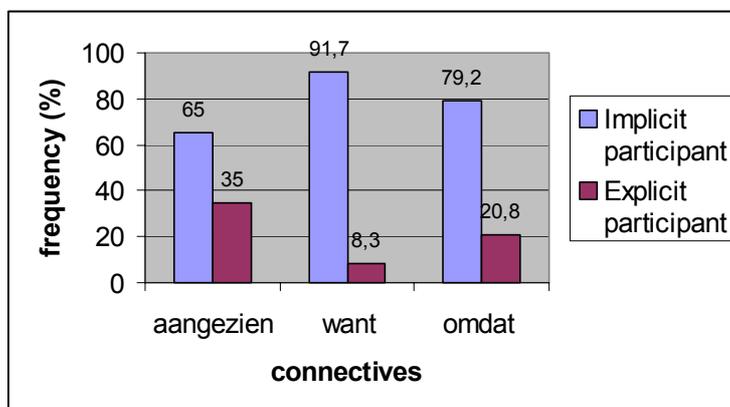


Figure 3: *Explicit or implicit presence of the causal participant co-occurring with aangezien, want, or omdat in epistemic relations (with participant)*

$$\chi^2 = 6.462, df = 2, p < 0.05 (p = .040)$$

$$want-aangezien: \chi^2 = 6.279, df = 1, p < .05 (p = .012)$$

The hypothesis that *want* encodes a higher SI level than *aangezien* is supported by two SI features: *want* more often has **first-person participants** (Figure 2) and these are more often left **implicit** in the first segment (Figure 3). If we consider these data together with those in Table 2, we believe we may conclude that the three Dutch causal connectives under investigation can be put on the SI scale in increasing order starting with *omdat*, followed by *aangezien* and *want* being highest. This is a striking result since the preliminary results for the French connectives placed *puisque*, supposedly the counterpart of *aangezien*, highest on the SI scale. So let us have a look at Figures 4 to 6 which display the SI features for the French connectives (again leaving out those cases which present no variation with respect to the causal participant involved, if at all).

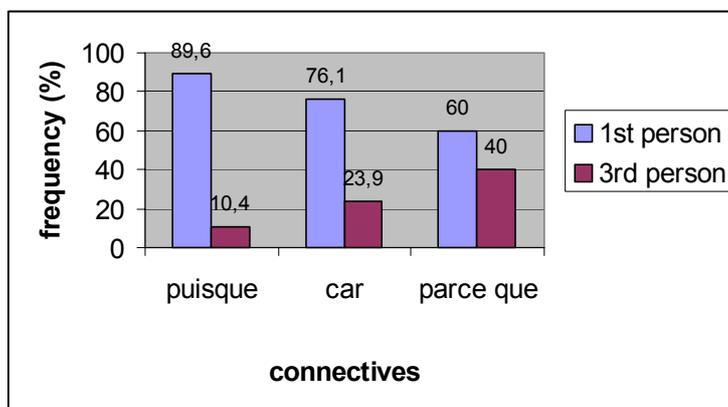


Figure 4: Identity of the causal participant co-occurring with *puisque*, *car*, or *parce que* in volitional and epistemic relations

$$\chi^2 = 14.542, df = 2, p < .001$$

puisque-car: $\chi^2 = 3.029, df = 1$, which is significant at $p = .041$ in a one-tailed test

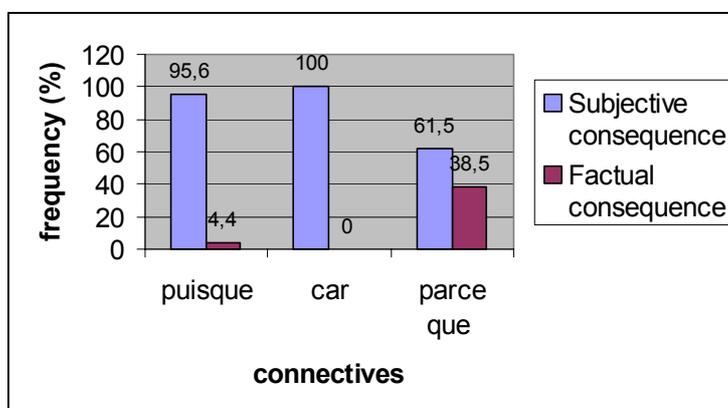


Figure 5: Type of conclusion co-occurring with *puisque*, *car*, or *parce que* in epistemic relations

$$\chi^2 = 2.258, df = 2, p > .05 (p = .323) \text{ N.S.}$$

puisque-car: $\chi^2 = 1.57, df = 1, p > .05 (p = .21) \text{ N.S.}$

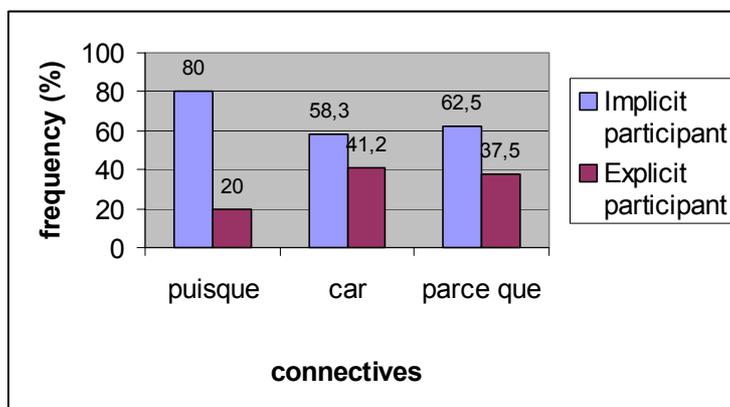


Figure 6: *Explicit or implicit presence of the causal participant co-occurring with puisque, car, or parce que in epistemic relations (if participant present)*

$$\chi^2 = 4.782, df = 2, p > .05 (p = .92) \text{ N.S}$$

$$\text{puisque-car: } \chi^2 = 4.5503, df = 1, p < .05 (p = .034)$$

Figures 4 and 6 confirm that in French *puisque* is higher on the SI scale than *car* and *parce que*, since *puisque* more often comes with implicit first person participants. Hence, in French the respective positions of the three connectives is different from that in Dutch.

In order to check whether these differences in distribution on the scale are the result of specific divergences in use of some of the connectives, or rather result from more general differences, we propose to make an inter-language comparison of the counterparts for each connective under investigation. The comparison will bear on the four SI aspects already accounted for. The results are given in Table 5 and Figures 7-8.

Relation	connective					
	<i>puisque</i>	<i>aangezien</i>	<i>car</i>	<i>want</i>	<i>parce que</i>	<i>omdat</i>
non-vol. causal		2	1	1	11	10
vol. causal	1	8	10	7	23	16
causal epis- temic	36	26	27	20	14	20
non-causal epistemic abductive	10	14	9	17	2	4
epistemic speech act	1			1		
	2		3	4		
	50	50	50	50	50	50

Table 5: *Inter-language comparison: relational interpretations*

puisque-aangezien : $\chi^2 = 12.72$, $df = 5$, $p < 0.05$

car-want : $\chi^2 = 5.18$, $df = 5$, $p = 0.39$ (N.S.)

parce que-omdat : $\chi^2 = 3.03$, $df = 3$, $p = 0.39$ (N.S.)

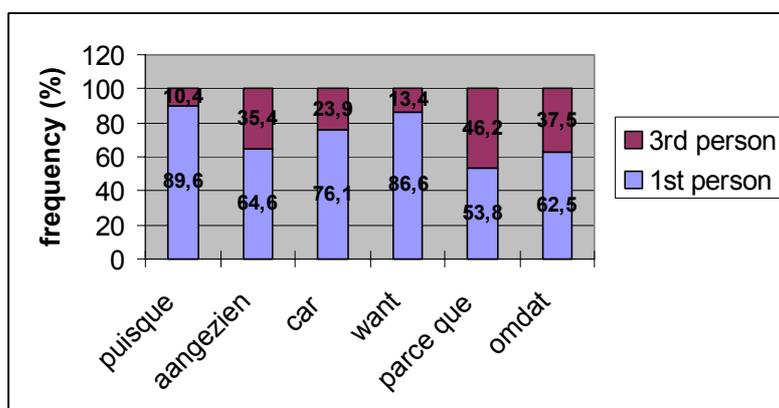


Figure 7: *Inter-language comparison: identity of the relational protagonist*

puisque-aangezien: $\chi^2 = 8.491$, $df = 1$, $p < 0.01$

car-want: $\chi^2 = 1.676$, $df = 1$, $p = 0.195$ (N.S.)

parce que-omdat: $\chi^2 = 0.608$, $df = 1$, $p = 0.436$ (N.S.)

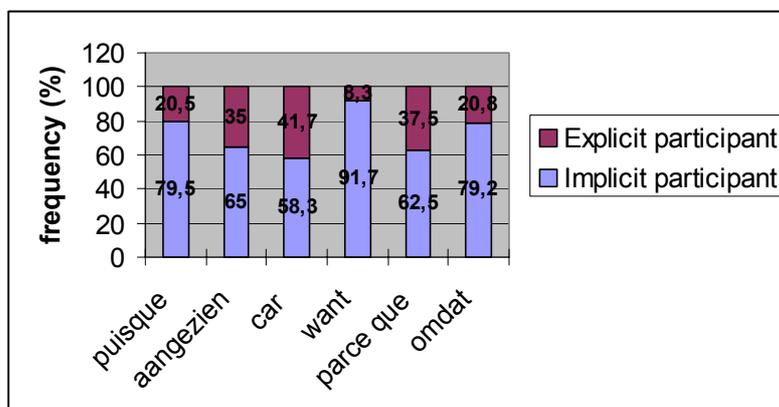


Figure 8: *Inter-language comparison: realisation of protagonist*
puisque-aangezien: $\chi^2 = 2.414$, $df = 1$, $p = 0.12$ (N.S.)
car-want: $\chi^2 = 9.023$, $df = 1$, $p < 0.01$ ($p = 0.003$)
parce que-omdat: $\chi^2 = 1.338$, $df = 1$, $p = 0.247$ (N.S.)

Table 5 and Figures 7-8 show that French *puisque* partially differs from Dutch *aangezien*. There are indeed significant differences between the two connectives in terms of relational interpretations (Table 5) and in terms of the identity of the causal participant (Figure 7). This leads us to the conclusion that *puisque* takes a higher position on the SI scale than *aangezien*. Furthermore, French *car* appears to be very close to Dutch *want*, the latter connective having, however, a higher SI since it occurs more often with implicit participants (Figure 8). Our data do not show any significant differences between the two low SI connectives *parce que* and *omdat*.

3.2. Speaker Involvement and connectives in use

In the previous section we have ordered a number of backward causal connectives on the speaker involvement scale. Now, what does it mean for a causal connective to be higher or lower on the Speaker Involvement scale? According to us, every connective brings some of its semantic content to the causal relation it is building. In addition to making explicit a causal relation between two segments, a connective also encodes a certain SI level which it contributes to the interpretation of its discourse environment. It is our claim that the proper SI level of the connective, i.e. its position on the SI scale, and the SI level of the given stretch of

discourse, i.e. the underlying causal relation, should concord. In other words, when the SI level of the connective is too high or too low to be combined with the level allowed for by the discourse environment, the use of the connective is inappropriate.

In the next three examples the differences in SI manifest themselves in the degree to which the speaker is presenting his stealing as justified. With *want* the speaker conveys the assumption that stealing from your boss is a perfectly natural thing to do for somebody in need of extra money for his family. With *omdat*, the speaker does not justify his behavior, but merely explains it: he presents a reason without tempting to generalize the validity of the relation between reason and action. The assumptions conveyed by *aangezien* are more like those conveyed by *want* than those conveyed by *omdat*. (Though *aangezien* may be somewhat weaker than *want*. Intuitions get rather subtle at this point).

- (18) a. Ik stal van mijn baas, *omdat* ik extra geld nodig had om mijn grote gezin te onderhouden.
 ‘I stole from my boss, [*omdat*] I needed extra money to support my large family.’
- b. Ik stal van mijn baas, *aangezien* ik extra geld nodig had om mijn grote gezin te onderhouden.
 ‘I stole from my boss, [*aangezien*] I needed extra money to support my large family.’
- c. Ik stal van mijn baas, *want* ik had extra geld nodig om mijn grote gezin te onderhouden.
 ‘I stole from my boss, [*want*] I needed extra money to support my large family.’

Another way of demonstrating differences in SI profiles is substituting the connectives for each other in contexts with a low degree of speaker involvement. For instance, when the speaker clearly distances himself from another protagonist by criticizing his actions, a low-SI connective should be more appropriate than a high-SI connective. Compare the next two examples, taken from the Dutch *omdat* corpus, in which *aangezien* appears inappropriate.

- (19) Afwezig waren vertegenwoordigers van de Nederlandse vakbeweging. “Dat is raar, maar ze willen niet *omdat*/**aangezien* het Europese werk voor hen een ver-van-mijn-bed-show is en ze ons een beetje zien als concurrenten”, aldus X.
 ‘The representatives of the Dutch trade unions were not there. “That is strange, but they don’t want to come [*omdat*/**aangezien*] they consider the European work as less important than their domestic affairs and they consider us as competitors”, says X.’
- (20) Er wordt met name in de sociale wetenschappen maar al te vaak verkrampt geschreven, *omdat*/**aangezien* men zich van tevoren tegen kritiek tracht in te dekken.
 ‘Especially in the social sciences we often encounter over-cautious writing, [*omdat*/**aangezien*] one tries to protect oneself from criticism in advance.’

Finally, a word should be said about the cross-linguistic comparison of the backward causal connectives. The data in section 3.1 clearly show that *puisque* and *aangezien* partially diverge from each other in that they do not naturally occur in the same types of discourse segments. In fact, our data with respect to the relational interpretations and the identity of the primary participant in the causal relation show that *puisque* has a higher SI than *aangezien*. As a result, *aangezien* is not used in the same way as *puisque* is. This might explain their diverging frequencies. In effect, the role of *puisque* seems to be largely taken over by *want* in Dutch, whose “official counterpart” *car* seems to be a little lower on the SI scale, *car* co-occurring less with implicit participants than *want*. At first view, *parce que* and *omdat* do occupy similar positions on the SI scale as appears from their occurrence in similar discourse environments.

A last question remains about the functional use of the connective *aangezien*. It is indeed fairly infrequent with respect to the other backward causal connectives, although not specifically bound to a restricted text genre (Degand 2000), and very close in meaning to *want*, although with a lower SI. A possible explanation could be syntactic in nature. Contrary to *want*, *aangezien* is a subordinating conjunction that can be used in preposed constructions like (21). Actually, these types of constructions are highly frequent in Dutch since nearly 50% of the

aangezien-segments occur in this ordering, as was shown in Table 1 above.

- (21) Aangezien we dit jaar niet op vakantie kunnen, zullen we maar veel eendagstripjes maken.
 '[Aangezien] we can't go on holiday this year, we will probably make a lot of one-day trips.'

A possible explanation for the high frequency of *aangezien* in preposed position could be that it is used instead of *want* when this latter connective is excluded for syntactic reasons, or at least reasons of information flow. Recent research (Degand 2000) has shown that the use of preposed vs. intermediate positioning in causal sequences is strongly linked to the preceding information. In particular, the first causal segment tends to pick up information contained by the previous theme or rheme. Since the connective *want* does not allow such syntactic manipulations and the connective *omdat* is too low on the SI scale; Dutch seems to have favored the use of preposed *aangezien* in such cases.

In French this phenomenon does not take place. The three connectives *parce que*, *car* and *puisque* differ significantly from one another. At the same time, substitution hardly ever leads to unacceptable sentences, but rather to semantic differences which we have explained in terms of SI.

4. Conclusion

In this paper, we have investigated whether an alternative account to the categorization of connectives, the Speaker Involvement scale, could be extended to backward causal connectives in Dutch and in French. To this end, we performed a corpus analysis of a number of backward causal connectives in terms of SI. In concrete terms this means first investigating the discursive context in which the connectives naturally occur and re-analyzing them in terms of a scale of increasing speaker involvement in the construction of the causal relation: from non-volitional to speech act relations via, in increasing SI-order volitional, causal epistemic and non-causal epistemic relations. The SI profiles thus developed were further supported by observations regarding the semantic effects of substituting a connective by one with a different SI level. According to

us, such a semantic analysis leads to fine-grained distinctions uncovering subtle meaning differences within a language but also cross-linguistically. In particular, it appeared that two supposedly translation equivalents *puisque* and *aangezien* do in fact not appear in the same contexts of use. According to our analyses, *puisque* encodes a higher level of SI than *aangezien* which should explain these divergences in use, as well as the high frequency difference of the two connectives in the two languages. Further research on bilingual and translated corpora is needed to show whether it is indeed the case that *puisque* is most often **not** translated by its “official” equivalent *aangezien*. The question remains open on which Dutch connectives fulfil the role of *puisque* in naturally occurring texts. According to our first observations, *want* should be frequent in these contexts but there might be other constructions playing a similar role.

An important outcome of this research is that connectives cannot be considered in isolation from one another. We believe to have shown that causal connectives, both in Dutch and in French, should be considered on a continuum from lowest to highest speaker involvement. The variation in the use of these connectives is due to the relative independence between connective meanings and categories of relational interpretations. This yields expressive possibilities for speakers who want to introduce assumptions in, or remove them from the interpretation of a certain relation. Such findings are of course important both for translation and for foreign language teaching. Finally, since the results we obtained are similar to those for forward causal connectives (Pander Maat & Degand 2001), we can conclude that the SI scale has proven to be valid for backward as well as forward causal connectives in Dutch and in French.

Acknowledgements

Liesbeth Degand is a research fellow at the Belgian Fund for Scientific Research (F.N.R.S.). At the time of this research, she was also supported by an I.U.A.P project from the Belgian Federal Government (I.U.A.P 4/19).

References

- Belle, W. van (1989). Want, omdat, en aangezien: een argumentatieve analyse. *Leuvense Bijdragen* 78, 435-556.
- Degand, L. (1998). Het ideationele gebruik van *want* en *omdat*: een geval van vrije variatie? *Nederlandse Taalkunde* 3, 309-326.
- Degand, L. (2000). Contextual constraints on causal sequencing in informational texts. *Functions of Language* 7, 173-201.
- Degand, L. (2001). *Form and Function of Causation. A theoretical and empirical investigation of causal constructions in Dutch*. [Studies op het gebied van de Nederlandse Taalkunde]. Peeters, Leuven.
- Degand, L. & H. Pander Maat (1999). Scaling causal relations in terms of Speaker Involvement. *Levels of Representation in Discourse, Working Notes of the International Workshop on Text Representation*, Edinburgh University, 45-54.
- Degand, L. & T. Sanders (1999). Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives. *Levels of Representation in Discourse, Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.
- Dijk, T.A. van (1979). Pragmatic connectives, *Journal of Pragmatics* 3, 447-456.
- Groupe λ -I (1975). Car, parce que, puisque. *Revue Romane* 10, 248-280.
- Halliday, M.A.K & R. Hasan (1976). *Cohesion in English*. Longman, London.
- Iordanskaja, L. (1993). Pour une description lexicographique des conjonctions du français contemporain. *Le Français moderne* 2, 159-190.
- Mann, W.C. & S.A. Thompson (1986). Relational propositions in discourse. *Discourse Processes* 9, 57-90.
- Martin, J.R. (1992). *English Texts: system and structure*. John Benjamins, Amsterdam.
- Moeschler, J. (1989). Pragmatic connectives, argumentative coherence and relevance. *Argumentation* 3, 321-339.
- Pander Maat, H. (1998). Classifying negative coherence relations on the basis of linguistic evidence. *Journal of Pragmatics* 30, 177-204.
- Pander Maat, H. & L. Degand (2001). Scaling causal relations and connectives in terms of Speaker Involvement. *Cognitive Linguistics* 12, 211-245.
- Pander Maat, H. & T. Sanders (1995). Nederlandse causale connectieven en het onderscheid tussen inhoudelijke en epistemische coherentie-relaties. *Leuvense Bijdragen* 84, 349-374.
- Sweetser, E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press, Cambridge.

Index

- Abraham, W., 167
adjunct, 27, 33, 37
Adriaens, G., 25
adverbial, 39
Afrikaans, 172
agentive, 106, 155, 156, 163, 164
Alphen, I. van, 9
ambiguity, 16
ambiguity avoidance, 168, 170
animacy, 98, 102, 107, 115
– of the pronoun object, 147
– of the subject, 154
ANS, 59, 60, 78, 114, 115, 122, 135
asymmetry between subject and object, 170
- Balk-Smit Duyzendkunst, F., 146
Barlow, M., 27
barrier (in Force Dynamics), 43, 69, 78, 79, 89
Beets, A., 62
Behaghel, O., 138, 159, 166
Behaghel's "Gesetz der wachsenden Glieder, 159
Belle, W. van, 183
benefactive, 27, 51, 52
beneficiary, 29, 30, 31, 32, 40, 41, 42
Bezooijen, R. van, 61
blending, 39
- Blom, A., 64, 65, 90, 93
Boogaart, P.C. uit den, 107
Booij, G.E., 148
Bree, C. van, 52
Brown, P., 68
Buitenhuis, P.A., 62, 63
Burrige, K., 166, 167, 172
Bybee, J.L., 24, 27, 54
- Caluwe, J. de, 7, 8
Caspers, J., 59
causal connectives, 175, 176, 180, 181, 182, 183, 186, 190, 194, 196, 197, 198
causality, 97, 101, 104, 108, 109, 128, 176, 177, 178, 186
coercive, 98, 99, 102, 105, 108, 109, 124, 128
direct, 97, 101
indirect, 97, 101
permissive, 98, 99, 102, 104, 105, 112, 116, 124
causative constructions, 43, 97, 99, 107, 119
analytic, 97
causeless constructions, 99, 106, 107, 108, 109, 110, 111, 112, 123
clitic, 140, 148, 149, 151, 170
cliticization, 148, 149
coherence, 28, 56, 80
coherence relations, 175, 176, 180, 183

- collocations, 3
 commands, 64, 68, 69, 77, 79, 83, 84, 85, 86, 88, 90, 91, 92
 Contini-Morava, E., 92
 contrastive analysis, 175, 183, 186
 contrasts in message fractions, 66
 conventionality, 33, 34, 37, 40, 44, 46, 52, 55
 conventionalized
 exploitations of a form, 62
 Cook, K., 61, 91
 Coppen, P.A., 91
 corpus analysis, 21, 28, 31, 36, 37, 40, 41, 45, 47, 48, 80, 99, 107, 112, 113, 122, 123, 124, 136, 138, 142, 145, 150, 156, 161, 166, 176, 181, 182, 183, 184, 185, 188, 195, 197, 198
 Croft, W., 98
 Cruse, D.A., 24
 Cruysse-van Antwerpen, I. van der, 69

 Deen, J.Y., 59, 61, 69
 default verb, 35, 46
 definiteness, 29, 30, 41
 Degand, L., 99, 100, 101, 175, 176, 177, 180, 183, 196, 197, 198
 Dewey, T., 172
 differentiation, 8, 16
 Dijk, T.A. van, 175
 Dik, S.C., 103, 114, 159, 166
 direct object, 29, 30, 31, 39, 40, 41, 51
 directive sentences and politeness, 69
 discourse, 91, 92, 175, 179, 180, 181, 185, 194, 196
 discourse markers, 61
 ditransitive, 51, 52, 53, 90
 drift, 142, 166, 167, 168, 169, 170, 171
 Duinhoven, A.M., 64, 86, 91, 93
 Dutch
 Middle Dutch, 86, 135, 141, 167, 169, 170
 Modern Dutch, 39, 59, 92, 103, 132, 133, 134, 135, 137, 148, 152, 155, 156, 158, 159, 163, 165, 167, 169, 170

 Early Immediate Constituents, 159, 161
 emergent properties, 3, 90
 encyclopedic knowledge, 18, 19
 endpoint, 42, 43, 44, 91
 English, 2, 3, 7, 16, 28, 32, 33, 34, 36, 38, 39, 40, 43, 46, 50, 51, 52, 59, 61, 62, 73, 80, 86, 90
 Modern English, 39, 171
 entrenchment, 24, 38, 54, 56
 epistemic, 175, 176, 177, 178, 179, 180, 181, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 197

 Fagan, S.M.B., 115
 focus, 124
 Foolen, A., 69, 78
 Force Dynamics, 43, 99
 form-meaning pairing, 33
 Fortuin, E., 93

- French, 5, 14, 15, 175, 176, 180, 181, 182, 183, 185, 186, 190, 192, 194, 197, 198
- frequency, 21, 31, 34, 47, 54, 88, 90, 181, 182, 196, 197, 198
- Geeraerts, D., 18
- gender, 8, 9, 10, 11, 12, 15, 17, 19, 20, 21, 22, 23, 24
- German, 2, 8, 17, 135, 140, 144, 163, 166, 167, 172
- Low German, 136, 172
- Gerritsen, M., 25
- Gestel, F.C. van, 135, 149
- given-new strategy, 85
- Givón, T., 86
- goal, 27, 42, 44, 46, 49
- Goddard, C., 15
- Goldberg, A.E., 28, 32, 38, 41, 50, 54, 90
- grammatical organization, 28
- grammatically determined word order, 166, 167, 172
- Gricean maxims, 62, 67
- Groupe λ -1, 183
- Haan, S. de, 65
- Haeseryn, W. et. al., *see* ANS
- Haiman, J., 86
- Halliday, M.A.K., 175
- Han, C., 91
- Hasan, R., 175
- Hawkins, J.A., 159, 161, 162, 166, 170
- Hebrew, 90
- hè-hoor* opposition, 61
- Heuven, V.J. van, 59, 61, 62, 66, 77
- Hoeksema, J., 90
- Hoekstra, T.A., 92
- Hopper, P., 27, 54
- Horst, J.M. van der, 134, 136, 139, 163, 164, 165, 171
- Huisman, J., 9
- Hyams, N.M., 90, 92
- hyponyms, 12, 13
- iconic placement of negator, 84
- iconicity, 186
- idiomaticity, 90
- degrees of, 90
- Ijbema, A., 67
- imperative structures, 59, 77, 78, 87, 89
- imperatives
- success imperative, 91
- Imperativity, 72, 73, 74, 76, 77, 78, 79
- inanimacy, 98, 102, 114, 115
- indefiniteness, 29, 30, 40, 41
- of the subject noun, 153
- indirect object, 27, 29, 34, 39, 51
- indirectness continuum, 105, 112, 124, 128
- infinitives, 67
- Infinitivus pro imperativo (IPI), 59, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92
- innovation, 86
- instrument, 38
- internet, 31, 47
- intransitivity, 43, 44
- invisible hand explanations, 167, 168, 169, 172

- Iordanskaja, L., 183
 Israel, M., 39
 Italian, 90
- Jackendoff, R., 32, 38, 54
 Jakobson, R., 12, 15, 19, 23
 Janssen, T., 69, 78
- Keller, 168
 Kemmer, S., 27, 43, 97, 99, 100, 101, 103, 104, 105, 106, 107, 109, 111, 115, 116, 120, 121, 125, 126, 127
 Kirsner, R.S., 59, 61, 62, 66, 69, 77, 79, 93, 138
 Knuttel, J.A.N., 62
 Koster, J., 146
 Kramer, A., 36, 41
- Lambooy, J.H.Th., 67
 Langacker, R.W., 18, 19, 23, 24, 27, 54, 61, 89, 90
 language change, 168
 language-independent preferred order of constituents, 159
 Lehmann, W.P., 148
 length difference, 159, 160, 161, 162
 Levinson, S.C., 68
 lexical strength, 24
 lexicon, 24, 54, 72, 90
 Loewenthal, J., 43, 97
 loss of morphological marking of subject vs. object, 168
 Lyons, J., 12, 13
- Mann, W.C., 175
 markedness, 10, 12, 13, 15, 17, 19, 23, 178
- Martin, J.R., 175
 maximalist vs. minimalist linguistics, 89
 memory, 34, 38, 50, 56
 long-term memory, 33
 middle constructions, 122, 127
 middle field, 135, 137, 141, 142, 154, 155, 163
 modal verbs, 86
 modality, 45
 Moeschler, J., 175
 morphology, 54, 67
 unitary analysis, 67
 movement, 32, 33, 38, 39, 48, 49, 178
 accompanying activity, 38, 39
- names for professions, 8, 9, 24, 25
 negation, 45, 83, 86, 90, 91, 92, 121
 network, 23, 24, 38, 51, 52, 53, 54, 55
 neutralization, 8
 Nieuwborg, E., 163
 Nijen Twilhaar, J., 135, 149
 nominal case marking, 133
 non-volitional, 175, 176, 178, 179, 186, 188, 197
- object preposing (of pronoun), 135
 object pronouns
 declination of frequency over time, 147
 frequency of object pronoun, 147
 oblique, 36, 37, 40, 42, 49
 opposition, 9

- equipollent opposition, 11, 14
privative opposition, 10, 11, 13, 14
order of pronominal objects and nominal subjects, 135
Osgood, C.E., 78
- Paardekooper, P.C., 64, 65
paired comparisons, 75
Pander Maat, H., 175, 176, 177, 180, 183, 198
particles, 59, 62, 68, 69, 70, 72, 77, 87, 88, 89
path, 28, 32, 33, 34, 35, 36, 37, 38, 39, 41, 42, 43, 44, 46, 48, 49, 50, 52
removing obstacles, 32
pitch, 77
polysemy, 13, 14, 15, 16, 17, 25, 39
possessive pronoun, 20, 29, 32
pragmatic particles, 77, 92
pragmatically determined word order, 166, 171, 172
pragmatics, 18, 59, 62, 69, 77, 79, 80, 86, 90, 91, 92, 175
preposition, 29, 37, 40, 42
productivity, 7, 8, 27, 52, 55, 91
Proeme, H., 64, 66, 93
profiling, 77
prohibitions, 83, 84, 86, 87, 88, 90, 91, 92
pronoun case, 144, 145
pronoun cliticization, 148
propositional content, 44, 177
prototypes, 36, 38, 47, 48, 49, 50
prototypical, 24, 180
psycholinguistics, 93
questionnaire, 59, 72, 75, 83, 88, 93
reflexivity, 29, 30, 31, 32, 34, 35, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 51, 116, 118, 119
full form, 31
strong form, 31
weak form, 31
Reichenbach, H., 91
repeated measures analysis, 75
resultative, 50, 51
Rinkel, T., 135, 149
- Sanders, T., 101, 176, 183
Santen, A. van, 7, 8
Sapir, E., 166, 170, 171
schema, 23, 24, 37, 38, 50, 55
schematic analyses, 61, 66, 67, 89
Schiffrin, D., 61
semantic role, 142, 147, 148, 151, 155, 156
semantic role of the subject, 155
semantics, 18, 43, 49, 52, 90, 92
Shannon, T.F., 131, 135, 136, 141, 145, 148, 149, 150, 152, 155, 156, 163, 164, 166, 168, 169, 172
Shetter, W.Z., 69
social knowledge, 18, 19
Source-of-Energy, 44
SOV to TVX/SVO, 171
Spanish, 90
speaker involvement, 176, 194, 195, 197, 198
Speaker Involvement (SI) scale, 175, 179, 180, 181, 185, 186,

- 187, 190, 192, 194, 196, 197, 198
- Speaker-Hearer relationship, 61
 profiling of, 63
- speech act, 44, 61, 68, 79, 175, 176, 177, 178, 180, 184, 185, 193, 197
- standard language, 52
- stereotypes, 17, 18
- STM, 59, 64, 65, 66, 67, 68, 69, 70, 71, 72, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 92
- STM-IPI opposition, 59, 93
- Stoett, F.A., 86
- Stukker, N., 101
- stylistic factors, 86
- subject prominence, 166, 167, 168
- subject type, 151
 common vs. proper, 151
- Suci, G.J., 78
- superordinate, 13, 15, 38, 50
- Sweetser, E., 62, 79, 175
- synergism vs. antagonism, 71, 79
- syntax, 29, 32, 33, 34, 38, 39, 54, 91, 182, 196, 197
- Talmy, L., 43, 98, 99, 112
- Tannenbaum, P.J., 78
- taxonomic network, 38
- Thompson, S.A., 166, 175
- topic prominence, 166
- topic-prominent constructions, 166, 167
- transitivity, 29, 33, 39, 43, 50, 51, 98
- Tuggy, D., 16
- types of pronouns, 140
- Ulenspieghel*, 135, 136, 137
- usage, 1, 19, 22, 23, 24, 25, 27, 28, 31, 38, 54, 56, 66, 68, 88, 93
- usualness, 73, 74, 75, 76, 78, 79, 80, 89
- utterance-final pragmatic particles, 59, 60, 78, 88
- vagueness, 15, 16, 25
- verb second in main clauses, 166
- verbs
 cognition verbs, 103, 104, 107, 124, 125, 127
 perception verbs, 103, 104, 123, 124
- Verhagen, A., 15, 27, 28, 34, 36, 43, 55, 97, 99, 100, 101, 103, 104, 105, 106, 107, 109, 111, 116, 125, 126, 127, 146
- Vismans, R., 69
- volitional, 44, 176, 177, 180, 181, 183, 184, 186, 187, 189, 191
- Wackernagel's Law, 148, 150, 166, 167
- Wal, M.J. van der, 171
- way construction, 28, 32, 33, 34, 36, 38, 39, 40, 44, 46, 48, 49, 50, 51, 52, 53
- Weerman, F., 135, 149
- Weijnen, A.A., 86, 171
- Weinreich, U., 78
- Wierzbicka, A., 18, 19
- word length, 159, 160, 161

Wouden, T. van der, 69, 73, 80

Yiddish, 172

Zaenen, A., 28

Zanuttini, R., 90

zich, 29, 30, 31, 33, 34, 35, 36,
37, 38, 39, 40, 41, 44, 45, 47,
49, 50, 52

zichzelf, 31, 32, 41