

University of Groningen

Topics in Corpus-Based Dutch Syntax

Beek, Leonoor Johanneke van der

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2005

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Beek, L. J. V. D. (2005). *Topics in Corpus-Based Dutch Syntax*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 3

Dative Alternations

In this chapter we investigate the different realizations of the Dutch dative alternation and the factors that influence this construction. For English, general linearization constraints have been claimed to influence the realization of ditransitive verbs. But in contrast to English, the syntactic category of the recipient argument and the order of the arguments may vary independently in Dutch. We thus expect that those linearization constraints influence only the order of the arguments, not the choice for a dative NP or a dative PP. This hypothesis is tested on the basis of quantitative data from automatically annotated corpora and focusing on pronominality and weight constraints. On the other hand we expect lexical preferences to target only the syntactic category of the recipient argument, not the relative order of the arguments. This is tested using the same techniques. The factors shown to influence the dative alternation are formalized in the framework of Optimality Theory. Finally, we provide additional evidence for our analysis of argument order variation based on a second construction, the *Accusativus cum Infinitivo*.

3.1 Introduction

The dative alternation is a very well-studied linguistic phenomenon and it is by no means specific to Dutch: much work has been done on the dative alternation in English, illustrated in (1) (Givón, 1984; Pinker, 1989; Levin, 1993; Goldberg, 1995; Wasow, 2002; Krifka, 2001; Levin and Rappaport Hovav, 2002; Bresnan and Nikitina, 2003; Bresnan et al., 2005, for example). Many different factors have been claimed to influence the dative alternation in English, ranging from general linearization principles, to semantics, lexical preferences and person effects.

- (1) a. Jo gave the student the book

- b. Jo gave the book to the student

In Dutch, the alternation is a bit more complex. In many aspects, the syntactic construction resembles the English situation: although we adopt the by now established term ‘dative alternation’ to refer to the different realizations of verbs with a recipient and a theme argument, no actual dative casemarking is involved, just as in English. Both NP and PP recipients take the non-subject form, which can only be distinguished from the subject form if the recipient is pronominal. However, in addition to the canonical double object construction and the dative PP construction in (2-a) and (2-b), Dutch has two more variants: both the double object construction and the PP construction occur with non-canonical word orders. In (2-c) we find the direct object shifted in fronted of the indirect object. In (2-d) we see that the ‘dative’ PP is shifted and precedes the direct object. Both variations violate the rule for canonical argument order $\text{SUBJ} < \text{OBJ2} < \text{OBJ1} < \text{OBL, XCOMP}$ for Dutch.¹

- (2) a. Vervolgens gaf hij mij geel.
 afterwards gave he me yellow
And then he gave me a yellow card.
- b. Vervolgens gaf hij geel aan de speler.
 afterwards gave he yellow to the player
And then he gave a yellow card to the player.
- c. Vervolgens gaf hij het mij.
 afterwards gave he it me
And then he gave it to me.
- d. Vervolgens gaf hij aan die speler een eerste officiële
 afterwards gave he to that player a first official
 waarschuwing.
 warning
And then he gave a first official warning to that player.

The existence of the non-canonical variants in Dutch teases apart two distinctions that are merged together in the English situation: where English has two variants which differ with respect to both the syntactic category and the

¹Throughout this chapter, we will use both ‘indirect object’ and OBJ2 to refer to the grammatical role to which the recipient argument is mapped. This is contrary to much work in Lexical Mapping Theory on English, where the recipient is assumed to map to OBJ1 in the double object construction. There is reason to assume that English and Dutch differ in this respect, e.g. Dutch does not generally allow the recipient of a ditransitive verb to be mapped onto the subject function in passive sentences. It does allow passive sentences with theme subjects and recipient objects.

order of the arguments,² Dutch has two separate alternations: one between NP and PP recipients and one between the canonical and non-canonical word order. This results in a total of four different realizations for ditransitive verbs, which give us the opportunity to study the two alternations separately. One could expect, then, that the factors that have been claimed to influence the dative alternation in English, now fall in two categories. A first category which influences the syntactic category of the recipient argument, for example a lexical preference of the verb for an NP or a PP recipient, and a second category which contains general constraints on argument ordering. In this chapter we investigate whether this expectation is borne out.

This chapter is structured as follows. We start with a brief discussion of the literature on the dative alternation in English, focusing on the different factors that have been claimed to influence the alternations (section 3.2). We then investigate to what extent these factors influence the Dutch dative alternation, and more specifically whether they influence the argument order alternations (sections 3.4 and 3.5) or the NP/PP alternation (section 3.6). We incorporate our findings in the framework of Optimality Theory (OT). The advantage of modeling the influencing factors in OT is that it allows for violable constraints and constraint interaction. As we will see, there are many constraints that are important for the dative construction, but most of them can be overridden. The constraint on canonical word order, for example, can be overridden in order to avoid a right aligned pronoun. This is difficult to model in other grammatical frameworks. We will see, however, that the dative alternation poses challenges for OT as well. Section 3.8 shows how the argument ordering constraints that we identified for the dative alternation can be applied to account for another case of argument reordering in Dutch, the AcI construction. Finally, in 3.9 we summarize our findings and conclude.

3.2 Previous Work

The literature on the dative alternation can be divided in two categories: analyses that focus on the order of the two complements, and analyses that focus on the category of the recipient argument. We start with a brief overview of the first category and continue in section 3.2.2 with analyses of the second category.

²Even for English, the inverted word order is occasionally found, mostly in biblical texts (i). We will not go into these exceptional examples here.

(i) ...and I will give to him a white stone, and on the stone a new name written, which no one knows but he that receives it. (Revelations 2:17)

3.2.1 Linearization Constraints

The issue of constituent ordering has been discussed extensively in the linguistic literature. This section does not aim to be an exhaustive discussion of the complete literature on this topic, but merely an overview of various approaches with some pointers to work within that approach.

One approach to linearization can build on a particularly long tradition: Behaghel already observed in 1909 that long and complex phrases tend to follow lighter material (Behaghel, 1909/10). Since then, this observation has been applied to various ordering phenomena, including the dative alternation Arnold et al. (2000); Wasow (2002); Hawkins (1994); Erteschik-Shir (1979). The general idea is that although the double NP construction is generally favored, heavy recipients may be realized as (right-aligned) PPs in order to avoid a violation of the general principle on word order that says that heavy constituents align right. The influence of weight on word order is widely accepted and rephrased for other languages, including German (Uszkoreit, 1987) and Dutch. The Dutch reference grammar (Haeseryn et al., 1997) formulates this influence in the Complexity Principle, which states that light constituents precede heavy constituents.

Almost as widespread is the idea that information structure influences word order. Gundel (1988) and Prince (1992) showed that the contrast between old and new information is in this respect similar to the contrast between light and heavy material: new information tends to follow old, topic information. In addition, Arnold et al. (2000) showed that although weight and givenness are not independent of each other, they do both have a distinct effect on word order. Haeseryn et al. (1997) rephrase this constraint as the Left-Right Principle, which is claimed to be the main principle responsible for ordering in the Dutch midfield. Uszkoreit (1987) did not formulate a similar principle for German, but instead broke up the influence of information structure in two separate constraints: pronouns precede full NPs and definites precede indefinites. As pronouns are by definition old information but NPs not, and as definites are often old, but indefinites usually new, it is not hard to see that the Left-Right Principle and Uszkoreit's constraints lead to similar results.

German and Dutch, though allowing for word order variation, do have a clear canonical word order that is generally preferred. This may be stated as a simple constraint, e.g. "subject precedes indirect object which precedes direct object" (Uszkoreit, 1987). The Dutch reference grammar has formulated the Inherence Principle. This principle says that whatever is more closely connected to the main verb, should be realized closer to the second pole (i.e. the right edge of the midfield, which is defined to be the position of the verb

cluster). As the theme is generally considered more closely connected to the verb than the recipient, this principle favors the direct object being realized closest to the second pole. This is in line with the canonical word order in double object constructions, but clashes with canonical word order in dative PP constructions. The Inherence Principle is related to the “natural constituent structure” (Vennemann, 1973). This principle, attributed to Renate Bartsch, states among other things that the closeness of constituents in the surface string reflects the hierarchical dependencies between them. More recently, Wasow (2002) formulated the concept of semantic connectedness. Wasow states that constituents that are closely connected semantically are very likely to appear in adjacent positions. He furthermore argues that this constraint may override the other constraints (weight, information structure) on word order.

Furthermore, ambiguity avoidance may influence linearization. Wasow (2002) tested its influence and concluded that the aim to avoid ambiguity has at most a minor effect on the ordering of constituents. In this paper we will see some data which may point to such an influence, but conclusive evidence is yet to be found.

In addition to these general linearization constraints, it is well known (e.g. Lenerz (1992)) that pronouns behave differently from lexical NPs with respect to word order. We already saw that pronouns generally precede full NPs (Uszkoreit, 1987). Two constructions in which this principle plays a crucial role are Wackernagel Movement (WM) in German and Object Shift (OS) in Scandinavian languages. WM involves the obligatory shift of certain types of pronouns in German to a left-peripheral position, preceding all other nominal arguments (except subjects, which may precede the pronoun). Müller (2001) accounts for WM making use of a personal pronoun scale. As we will see, a similar scale exists in Dutch. Several Scandinavian languages show a similar flexibility in the midfield, allowing reduced pronouns to ‘shift’ leftwards, as long as they do not precede the verb they are an argument of (Diderichsen, 1946; Börjars et al., 2003). For a unified account of WM and OS, see Thráinsson (2001).

A radically different approach on word order is taken by Reinhart (1996), who formulated a focus-driven approach to word order. She argues that the sentence focus depends on the position of the main stress: the focus of IP is a(ny) constituent containing the main stress of IP. Usually, main stress falls on the right edge of the middle field in Dutch. If the focus of IP is a constituent that does *not* contain the rightmost phrase in the middle field, focus and stress do not coincide naturally. There are two ways to fix this: either the stress shifts to another position or the order of the constituents is changed through scrambling. Reinhart claims that scrambling is more

economical than stress shift and therefore the preferred strategy for stress (and thus focus) assignment.

The non-canonical versions of the double object construction and the dative PP-construction could be regarded as a form of scrambling and one could argue that even the NP/PP alternation could be explained in the same manner, as the canonical argument order for the dative PP and the double object construction do differ. Such a focus based approach would nicely explain why the phonologically weak pronoun *het* almost always shifts. After all, this pronoun is never stressed. It would also account for the fact that we find very few emphasized forms of the pronouns (e.g. *hijzelf*, ‘he himself’) in shifted position: the emphasized forms are focused and thus should stay at the right edge of the middle field in the canonical stress position. However, the data in (3) (grammaticality judgments hers), on which Reinhart (1996) bases her theory, are controversial, to say the least.

- (3) a. *Ik heb de krant nog niet gelezen, maar ik heb het boek
I have the newspaper yet not read but I have the book
al wel gelezen.
already indeed read
- b. Ik heb nog niet de krant gelezen, maar ik heb al
I have yet not the newspaper read but I have already
wel het boek gelezen.
indeed the book read
I haven't read the newspaper yet, but I did read the book already.

In any case, focus cannot be the full explanation for the non-canonical argument order: the alternation between canonical (4-a) and non-canonical (4-b) orderings persists even if both arguments are phonologically weak pronouns and therefore necessarily unstressed (4). In what follows, we do not discuss the role of stress and focus in the dative alternation, but instead focus on lexical and syntactic features.

- (4) a. “Blijf liggen”, zei hij, “ik geef je het wel.”
stay lying said he I give you_{weak} it_{weak} so
“Stay down”, he said, “I’ll hand it to you”.
- b. Ja ’k zal ’t ’m zeggen.
yes I will it_{weak} him_{weak} say
Yes, I’ll tell him that.

We have discussed four factors which may influence the order of the arguments in the dative alternation: canonical word order, weight, information

structure/focus, pronominality and ambiguity avoidance. In the research reported on below, information structure and ambiguity avoidance are briefly mentioned when applicable, but the main focus is on canonical word order, weight and pronominality.

3.2.2 The NP/PP alternation in English

As argument order and recipient realization are inseparable in the dative alternation in English, analyses of the construction have employed both the difference in ordering and the difference in syntactic category. In the previous subsection, we highlighted some approaches that focus on general ordering principles. We now turn to some analyses of the dative alternation that focus on the NP/PP alternation specifically.

Levin (1993) argues that verbs can be classified according to their meaning. Within these classes, all verbs are assumed to show similar syntactic behavior, but across classes, their behavior may vary. For instance, *send verbs* (e.g. ‘hand’, ‘mail’) and *transfer of message verbs* (e.g. ‘ask’, ‘read’) are assumed to alternate, while *manner of speaking verbs* (e.g. ‘whisper’, ‘shout’) and *bill verbs* (e.g. ‘bill’, ‘fine’) are not: these verbs supposedly only occur with the PP and NP alternant respectively. Lapata (1999) tested the empirical value of the semantic verb classes described by Levin (1993) with corpus based methods and concluded that there are statistically significant differences in the frequencies with which certain verbs occur in NP and PP ditransitive constructions. On the other hand, Bresnan and Nikitina (2003) convincingly showed that these are mere tendencies, rather than categorical differences.

Krifka (2001) and Pinker (1989), among others, take this one step further. Building on the work by Green (1974), they adopt a classification of verbs in verb classes, and argue that the reason that certain classes of verbs are incompatible with one of the two realizations is that the variants have different meanings and the verb meaning may be incompatible with one of them. The assumed meaning for the PP construction can be paraphrased as ‘x causes y to go to z’ and the meaning of the double NP construction as ‘x cause z to have y’. According to this line of explanation, there is no dative alternation proper: the double object construction and the PP construction are not alternative ways of expressing the same meaning, but they are realizations of different meanings. Bresnan and Nikitina (2003) provide examples of alternating dative syntax in contexts of repetition, which form a challenge for this approach. The view that the dative alternation reflects a difference in meaning contrasts with the widespread view that the verb means the same in both constructions, whether the two are related via syntactic transform-

ations (Larson, 1988; Aoun and Li, 1989) or different argument expressions of the same verb (Butt et al., 1997). Monosemic analyses of the dative alternation only explain the existence of the two structures, not the choice for one of them in a specific situation. They either predict free variation or require additional mechanisms such as general ordering or lexical constraints to determine the particular realization of a ditransitive verb.

Besides lexical preferences and semantic differences, there are also accounts of the NP/PP alternation (partly) based on structural constraints on the realization of the dative argument. Bresnan and Nikitina (2003) argued that local (1st or 2nd person) recipient arguments prefer to be realized as objects, not obliques, while nouns prefer the dative PP structure. This is formalized in the violable constraint conjunction $\text{HARMONY}(1,2): *NP_{Noun} \& *PP_{1,2Person}$. As a result of $\text{HARMONY}(1,2)$, local recipients will generally lead to double object constructions instead of dative PP constructions. Bresnan et al. (2005) furthermore report that the animacy of the recipient heavily influences its syntactic category (with inanimate recipients strongly preferring the dative PP structure).

This research will touch on the influence of person and animacy, but it will focus on verbal preferences for a PP or a dative NP structure. In English, these preferences will influence both the syntactic category of the recipient argument and the relative order of the theme and the recipient, because these two factors are inseparable. The hypothesis is that in Dutch, where syntactic category and order may vary fairly independently of each other, these lexical preferences influence only the choice for an NP or PP, not the argument ordering. This hypothesis is tested on the basis of frequency information from corpus data. The influence of these verbal preferences is then incorporated in our OT model of the dative alternation.

3.3 Preliminaries

In this section we describe the lexical resources that we used and the methodology that we applied for the corpus-based research described in the following sections of this chapter.

3.3.1 Resources and methodology

Corpora contain valuable information about the distribution of different realizations of the dative construction. A potential problem is that the structures we are interested in (the four alternants of the Dutch dative alternation) are

specific and complex syntactic structures, which cannot be retrieved from corpora on the basis of simple pattern recognition. Therefore, we used syntactically annotated and automatically parsed data in our corpus study. Two annotated corpora were used, the annotated part of the Corpus of Spoken Dutch (CGN, about 1M words (Levelt, 1998)) and the Alpino Treebank (the annotated CDB newspaper part of the Eindhoven Corpus, about 150K words (van der Beek et al., 2002a)), which are both annotated with dependency structures (Moortgat et al., 2001).

When the annotated corpora proved too small for statistically relevant results, we used a corpus of over 75M words of newspaper text (Twente Nieuwscorpus) that was automatically parsed by the Alpino parser (Bouma et al., 2001; van der Beek et al., 2002b). The parser outputs the same dependency structures as those used in the annotated corpora. With a 85.5% parsing accuracy (measured over the dependency relations), the quality of the annotation in the automatically parsed corpus is lower than the manually annotated corpora—although 100% accuracy is never reached, not even in hand-annotated corpora. The biggest problem with respect to the performance of the grammar is to make sure that the data is found, given that it is in the corpus. How do we know that the grammar will not systematically misparse certain double object constructions? The chances that the grammar does not recognize a dative construction are small. The hand-written Alpino grammar overgenerates, putting very few constraints on the argument order in double object constructions and always building a dative parse if possible. However, from all parses that are generated, only one is included in the corpus: the most probable parse is selected by a statistical Maximum Entropy model. We cannot look into this model to see what preferences it has deduced from manually annotated data. Here, we must rely on manual inspection of the data, to see if 1) the extracted candidates are true datives and 2) to see if any obvious constructions are missing. We did not find any signs of systematic errors in the data. As an extra precaution, we used queries that abstracted away from the argument order when extracting dative constructions from automatically parsed data (including both orders and sorting them manually), so that ordering errors by the parser did not influence the data collection.

The syntactically annotated corpora were queried using `dt_search` (Bouma and Kloosterman, 2002), a tool which allows us to query the treebank on dependency relations, syntactic category or part-of-speech and linear order. See chapter 1 for more general information on Alpino or `dt_search`.

We excluded from our search all instances of (in)direct object topicalization, all (wh)relativizer direct and indirect objects and all clausal objects such as that-clauses because in these sentences, the order of the arguments is

	NP NP _{unshift}	NP NP _{shift}	NP PP	PP NP	TOTAL
CGN	226	33	63	8	334
Alpino	122	7	43	10	182

Table 3.1: Distribution of the three alternants of the dative alternation Dutch manually annotated corpora.

determined by other factors. Passive sentences and instances of the *krijgen*-passive (the “*get*-passive”, a dative passive construction) were also excluded. The motivation for this is that the direct object (in the regular passive) or the indirect object (in the *krijgen*-passive) surfaces as the subject of the matrix clause, therefore the word order rules for subjects applies here. Finally, we excluded all instances of *er*-recipients. In these sentences, illustrated in examples (5), the recipient argument is third person, inanimate and singular and realized as a pronoun inside a PP. In these cases, so-called R-pronouns (*er/daar* ‘there’, *hier* ‘here’) are used instead of the regular third person neuter singular *het* ‘it’ and this pronoun is often fronted. The preposition stays in position, resulting in a split PP. The alignment of *er* is a characteristic of R-pronouns, not a characteristic of the dative construction.

- (5) Ik geef daar geen les aan.
 I give there no class to
I won't teach those.

3.4 Linearization: the Double Object Construction

In both the double object construction and the dative PP construction, the canonical word orders (NP NP_{unshift} and NP PP) are much more frequent than the shifted alternants (see fig. 3.1). Furthermore, the double object construction is much more frequent than the dative PP, and the canonical word orders (NP NP_{unshift} and NP PP) are much more frequent than the shifted alternants.

Although the general distribution is highly skewed, the chances of finding one of the less frequent realizations increases considerably under certain conditions. The following sections investigate what those conditions are. We first look at general linearization factors that influence the order of the objects in the double object construction.

	NP	NP _{unshift}	NP	NP _{shift}	NP PP	PP NP	TOTAL	
CGN		143		33		57	3	247
Alpino		45		6		21	3	83

Table 3.2: Distribution of dative alternation realizations with one word themes in manually annotated corpora.

3.4.1 Pronominality

The general distribution of the dative alternation changes drastically if we restrict the object to one lexical item only (we do allow additional function words such as determiners). While the numbers for the shifted double object construction hardly changed in table 3.2, the numbers for the unshifted and PP variants dropped by 10-70%. This may look like a strong influence of weight on constituent ordering, but it is actually caused by the fact that a shifted direct object almost exclusively occurs with pronominal direct objects.

Of all shifted direct objects in our manually annotated data, only one example contained a shifted full NP (6). In this example, we find the archaic dative marking on the indirect object. We assume that it is this overt dative marking that makes available the freer word-order and that Direct Object Shift (DOS) is generally restricted to pronouns.³

- (6) [daar] heeft Paul Badura-Skoda het nieuwe pianoconcert van Frank
 there has Paul Badura-Skoda the new piano_concert of Frank
 Martin den muzikale volke voorgesteld.
 Martin the_{dat} musical_{dat} people_{dat} presented
there, Paul Badura-Skoda presented the Frank Martin's new piano concert to the musical people.

With its restriction to pronominal objects, the Dutch object shift resembles the Scandinavian Object Shift (OS) and Wackernagel Movement (WM) in

³However, Zwart (1997) presents examples that show that NP-DOS with definite NPs is not impossible:

- (i) dat Jan het boek Marie terug gegeven heeft.
 that Jan the book Marie back given has
that Jan gave the book back to Marie.

No examples of this kind were found in our corpora. We suspect the exceptional definite NP shift to be a focus effect and leave this and other effects of focus on word order for future research.

German. But there are several differences: Scandinavian OS and WM apply to both direct and indirect objects, and may involve shifting beyond the subject. As a result, we find (in)direct objects preceding the subject in those languages. In Dutch, indirect objects may shift in front of the subject in some cases, although this is rare in ditransitive sentences. Direct objects on the other hand never precede subjects unless via topicalization. Furthermore, the Scandinavian shift depends on the position on the verb. This is not the case in Dutch.

English also has a constraint on the distribution on pronouns in the dative alternation: *NP PRO (Bresnan and Nikitina, 2003; Erteschik-Shir, 1979; Collins, 1995). This constraint states that personal pronouns, but not demonstratives or indefinite pronouns, are avoided when following full NPs if both are objects. The Dutch situation differs from English in that shift also occurs with demonstrative pronouns as well as personal pronouns, and that *het* ‘it’ not only precedes NPs but also other personal pronouns.

Not all direct object pronouns shift always. While the pronoun *het* ‘it’ usually shifts irrespective of the category of the indirect object, most other personal pronouns and the demonstratives shift if the indirect object is a full NP (7-a)⁴, but stay in their canonical position if the indirect object is a personal pronoun (7-b). First and second person pronouns do not shift. Not only are there no shifted local pronouns in the corpus, but made up examples of local pronoun DOS also lead to ungrammaticality under the intended reading (indicated with a %; example (8-b) is grammatical under the reading without DOS, i.e. the reading with a recipient *jou* ‘you’).

- (7) a. Ja, vertel dat de buurvrouw maar.
 yes tell that the neighbour DISC-PART
Yes, go ahead, tell it to the neighbour.
- b. Heeft hij je dat niet verteld?
 Has he you that not told
Didn't he tell you that?
- c. 'K zal 't hem zeggen.
 I will it him tell
I'll tell it to him
- (8) a. De student wijst 'm de docent aan.
 the student points him the teacher at
The student points him out to the teacher.
- b. %De student wijst jou de docent aan.
 the student points jou the teacher at

⁴DISC-PART indicates a discourse particle.

Shifted		Canonical	
542	het (it)	372	dat (that)
45	dat (that)	83	dit (this)
21	't (<i>it_{reduced}</i>)	51	het (it)
19	ze (them)	28	die (that)
7	dit (this)	24	hem (him/it)
4	u ⁵ (<i>you_{honorific}</i>)	14	zich (himself/herself)
4	hem (him/it)	8	hetzelfde (<i>it_{same}</i>)
4	die (that)	4	me (me)

Table 3.3: Direct object pronouns in constructions with two pronominal objects

The student points you out to the teacher.

Table 3.3 shows the most frequent direct object pronouns in double object constructions where both arguments are pronominal. The data are based on the automatically parsed TwNC corpus. The frequency lists confirm the intuition that *het* shifts while demonstratives usually do not shift in front of another pronoun. Importantly, the table shows that the distinctions are not categorical: we do find *het* (it) in the canonical object position, although ten times less frequently than in the shifted position. Manual inspection showed that these unshifted *het* objects are not (all) the result of parse errors. The one place where we would not expect any variation is with the local pronouns, as even made up examples were ungrammatical. It is therefore surprising that we do find four occurrences of *u* (*you_{honorific}*). Further inspection showed that these are the result of parse errors, however.

The pattern of pronouns preceding full NPs, and *het* ‘it’ preceding both pronouns (demonstratives, animate pronouns) and full NPs resembles the differentiation of pronouns proposed for German by Müller (2001) to account for Wackernagel movement (9). He argues that if a certain pronoun ‘moves’ under certain conditions, then all weaker pronouns do, too. In the Dutch dative alternation, there are two different conditions: either the indirect object is a full NP or it is a pronoun. In the first case, anything weaker than a local pronoun shifts. In the second case, usually only *het* ‘it’ shifts.

- (9) $\text{Pron}_{strong} > \text{Pron}_{unstressed} > \text{Pron}_{weak} > \text{Pron}_{reduced} > (\text{Pron}_{clitic})$
 +stress +anim -anim es ‘it’ ’s

⁵All occurrences of the local pronoun *u* result from parse errors.

We describe this pattern with a set of linearization constraints. All constraints take the form $\text{PRO}_x < y$, where x indicates a type of pronoun and y indicates some other type of argument nominal.⁶ The pronoun scale differentiates between animate and inanimate pronouns. Although the dative alternation does not provide evidence for this differentiation, we formulated separate constraints. We thus predict to find differences in the distributions of the two pronouns. In section 3.8 where we treat the Dutch AcI-construction, we will see that this expectation is borne out.

- (10) $\text{PRO}_{it} < \text{NP}/\text{PRO}$: the pronoun *het* precedes NPs and pronouns in the midfield.

$\text{PRO}_{3rd/inanim} < \text{NP}$: inanimate personal and demonstrative pronouns precede full NP arguments in the midfield.

$\text{PRO}_{3rd/anim} < \text{NP}$: animate personal and demonstrative pronouns precede full NP arguments in the midfield.

$\text{PRO}_{local} < \text{NP}$: local pronouns precede full NP arguments in the midfield.

The constraints on the linearization of pronouns are in competition with the constraint on canonical word order (11): a pronominal direct object violates $\text{PRO} < \text{NP}$ if it follows the NP indirect object, and it violates the canonical word order constraint if it precedes the indirect object. From the non-canonical example sentences (7) we can conclude that some of the constraints on pronoun linearization outrank the canonical word order constraint, but the ungrammaticality of (8) shows that not all pronoun constraints do. The interaction of the constraints is illustrated in tableau 3.4. We merged the linearization constraint on animate and inanimate personal/demonstrative pronouns for clarity, and included references to (made up) example sentences. It is important to note that in these examples, a star indicates a candidate that is *categorically ungrammatical*, while a question mark indicates a *significantly less frequent* candidate.

- (11) $\text{O2} < \text{O1} < \text{OBL}$: OBJ2 precedes OBJ1 precedes OBL

⁶We merged the two constraints $\text{PRO}_{it} < \text{NP}$ and $\text{PRO}_{it} < \text{PRO}$ for clarity.

Input: <i>gives</i> (<SUBJ><OBJ1><OBJ2>)		$Pro_{it} < NP/PRO$	$Pro_{3rd} < NP$	$O2 < O1 < OBL$	$Pro_{loc} < NP$	
OBJ1='the book' OBJ2='de student'	☞ NP NP _{unshift} NP NP _{shift}			*!		(12-a) (12-b)
OBJ1='it' OBJ2='de student'	☞ NP NP _{unshift} NP NP _{shift}	*!		*		(12-c) (12-d)
OBJ1='it' OBJ2='him'	☞ NP NP _{unshift} NP NP _{shift}	*!		*		(12-e) (12-f)
OBJ1='that' OBJ2='the student'	☞ NP NP _{unshift} NP NP _{shift}		*!		*	(12-g) (12-h)
OBJ1='that' OBJ2='him'	☞ NP NP _{unshift} NP NP _{shift}			*!		(12-i) (12-j)
OBJ1='you' OBJ2='the student'	☞ NP NP _{unshift} NP NP _{shift}			*!	*	(12-k) (12-l)

Table 3.4: Shifted vs. canonical double object constructions

- (12)
- a. Bo geeft de student het boek.
Bo gives the student the book
 - b. *Bo geeft het boek de student.
Bo gives the book the student
 - c. ?Bo geeft de student het.
Bo gives the student it
 - d. Bo geeft het de student.
Bo gives it the student
 - e. ?Bo geeft hem het.
Bo gives him it
 - f. Bo geeft het hem.
Bo gives it him
 - g. ?Bo geeft de student dat.
Bo gives the student that
 - h. Bo geeft dat de student.
Bo gives that the student
 - i. Bo geeft hem dat.
Bo gives him that
 - j. ?Bo geeft dat hem.
Bo gives that him
 - k. Bo raadt de student_{io} jou_{do} aan.
Bo recommends the student_{io} you_{do} PART
 - l. *Bo raadt jou_{io} de student_{do} aan.
Bo recommends you_{io} the student_{do} PART

The tableau shows how the constraints interact to account for the most frequent patterns in various types of sentences. Pronominal direct objects will shift if the indirect object is a full NP, in order to avoid a violation of the constraint on the linearization of pronouns, which is higher ranked than $O2 < O1 < OBL$. If both objects are pronominal (but not *het* or local), $O2 < O1 < OBL$ is the highest ranked constraint that is violated by one argument order but not the other and thus determines the optimal candidate: OBJ2 precedes OBJ1. *Het*, on the other hand, will shift no matter what the category of the indirect object is, because a violation of $PRO_{it} < NP/PRO$ is worse than any other right aligned pronoun or a non-canonical word order. $PRO_{local} < NP$ is outranked by $O2 < O1 < OBL$, preventing local pronouns from shifting.

Our findings contradict the claim in Zwart (1996) that only what he calls ‘reduced’⁷ direct object pronouns can shift: the demonstratives were

⁷Zwart’s notion of ‘reduced’ does not correspond to weak pronouns in the personal

among the most frequently shifted pronouns and we also found non-reduced examples of third person pronouns. We do see a tendency, though, of the third person reduced pronouns *'m* (him, it) and *ze* (them) to group with *het* if the antecedent is inanimate. In this case, they tend to shift, even if the indirect object is a pronoun. We do not have enough data for a quantitative evaluation of this intuition, but integration of it in our model is straightforward if it proves correct: they would be grouped together with *het*. In any case, it does not seem very likely that *all* reduced pronouns shift: objects that take the form of reduced local pronouns (*me* ‘me’, *je* ‘you’) were never found preceding the indirect object.

3.4.2 Gradient patterns

The tableau shows the interaction for the most frequent pattern for each combination of direct object and indirect object, but we have seen that the preferences are not categorical. Classic OT cannot account for these less frequent patterns, as it assumes a strict ranking of constraints (13). Given this strict ranking, C_3 can never dominate C_1 or C_2 , and alternative outcomes are ruled out. Variation is only possible in classic OT if the constraints on which two candidates differ are equally strong (i.e. in a stratum, see also table 1.5 in section 1.3.2). In this case, free variation is predicted. However, we have seen that one pattern may be much more frequent than other patterns: we do not find the fifty-fifty distribution that we would expect if it were free variation.

- (13) Strict constraint ranking in classic OT
 $C_1 \gg C_2 \gg C_3$

We would need the stochastic OT implementation of Boersma and Hayes (2001) to account for the alternative patterns (see chapter 1 for a brief introduction in Stochastic OT). Boersma and Hayes assume that the constraints are ranked on a linear scale, with higher values corresponding to stronger constraints, and that they are evaluated stochastically. Whenever a candidate set is evaluated, the exact position of a constraint on the scale is determined. This exact numeric value depends on its ranking, but is perturbed by a random variable, which models the noise in the system. For two constraints $C_1 \gg C_2$ that are ranked closely together, it is possible that because of this noise, the actual selection point of C_2 is sometimes higher than for C_1 , leading to an alternative ranking and a different optimal output.

pronoun hierarchy (Müller, 2001) above. Zwart (1996) distinguishes between different forms of the pronoun, not between properties such as person or animacy.

Returning to argument order alternations, we find both categorical distinctions and gradient patterns in the DOS. As local pronouns never shift, the distance between the canonical word order constraint and the constraint on local pronouns must be large enough for the reversed order to be practically impossible. On the other hand, we do find infrequent occurrences of the canonical word order with a 3rd person direct object pronoun and an NP indirect object (14): on a total of 139 occurrences of this combination in the TwNC corpus, 19 were in the canonical word order.

- (14) Als je de patiënt dat kunt besparen, moet je dat doen.
 if you the patient that can save must you that do
If can can spare this to the patient, you should.

Thus, the distance between the canonical word order constraint and $\text{PRON}_{3rd} < \text{NP}$ must be smaller, allowing for a chance that canonical word order outranks linearization constraints in some evaluation.⁸

3.4.3 Weight

Under our present analysis, the DOS is solely driven by the syntactic category of the objects: NPs, personal or demonstrative pronouns or *het*. But pronominality is not independent of syntactic weight: pronouns are the lightest possible NPs. Thus, the pronominal DOS is in line with the Complexity Principle and Uszkoreit's weight principle. But we did not differentiate between heavy NP recipients and light NP recipients, although the weight principle would predict the former to allow DOS more easily than the latter. Table 3.5 lists the average weight of the direct and indirect object in all four variants of the dative alternation, as well as the OBJ1/OBJ2 weight ratios.⁹ The weight is simply expressed as the number of words (discarding the preposition in the PPs). More sophisticated definitions of weight could be applied, such as number of nodes in a syntactic tree. However, it has been shown that different formulations of syntactic weight all lead to similar results (Wasow, 2002; Szmrecsányi, 2004). We see that the average weight of the indirect object in shifted double NP constructions (1.09 and 1.71) is *lower* than in the canonical double object construction (1.40 and 2.43), con-

⁸We also found non-canonical examples of demonstrative pronouns with personal pronouns indirect objects: 18 of 506 occurrences. To account for this word order, we need to assume a constraint $\text{PRO}_{3rd} < \text{PRO}$, similar as for the pronoun *het* (see also footnote 6), which is usually outranked by the constraint on canonical word order. With STOT, there will be some (small) chance on this constraint outranking canonical word order.

⁹We controlled for extraposition by only including sentences in which recipient and them were followed by the verbal cluster which indicates the right edge of the middle field.

				OBJ1	OBJ2	OBJ1/OBJ2
CGN	NP	NP _{unshift}	(N=231)	3.75	1.40	2.68
Alpino	NP	NP _{unshift}	(N=123)	5.87	2.43	2.42
CGN	NP	NP _{shift}	(N=33)	1.03	1.09	0.94
Alpino	NP	NP _{shift}	(N=7)	1.71	1.71	1.00
CGN	NP	PP	(N=56)	2.02	1.93	1.05
Alpino	NP	PP	(N=17)	2.53	2.29	1.10
CGN	PP	NP	(N=7)	4.71	2.71	1.73
Alpino	PP	NP	(N=8)	3.50	3.25	1.08

Table 3.5: Average weight per grammatical role in number of words.

trary to what the Complexity Principle would predict. If we control for a pronominality effect and exclude pronominal recipients, the average recipient weight increases and the differences between the two alternants get smaller. But with averages of 2.35 (Alpino) and 1.68 (CGN) for the canonical double object construction and 3.18 (Alpino) and 1.93 (CGN) for the canonical dative PP construction, the PPs are still heavier than the NPs. We conclude that syntactic weight does not have the expected influence on the DOS. This is surprising, as weight is generally assumed to influence linearization via the principle ‘light precedes heavy’.

We can conclude that the argument order in the double object construction is influenced by one important constraint on word order in general, namely the principle that states that pronouns precede full NPs, which in turn is related to the principle that old information precedes new information. The other important linearization principle, light constituents precede heavy constituents, could not be shown to influence the ordering of the two objects in the expected way.

3.5 Linearization: Dative PP Shift

We now turn to the second ordering alternation in the dative construction in Dutch: the argument order alternation in the recipient PP construction. This dative PP is most often realized in its canonical position following the direct object (15-a), but can also be found preceding the direct object (15-b), where it violates the principle of canonical word order.

Dataset			OBJ1	OBJ2	OBJ1/OBJ2
CGN	NP PP	(N=63)	1.62	2.57	0.63
Alpino	NP PP	(N=43)	3.70	5.21	0.71
CGN	PP NP	(N=8)	5.63	1.63	3.45
Alpino	PP NP	(N=10)	3.00	3.30	1.45

Table 3.6: Average number of words of PP recipients.

- (15) a. Als de speaker die treffer abusievelijk aan Amokachi
 when the speaker that hit mistakenly to Amokachi
 toekent, grijpt hulptrainer Jo Bonfrère in.
 assigns intervenes assistant-coach Jo Bonfrère PART
*When the speaker mistakenly assigns the goal to Amokachi, the
 assistant-coach intervenes.*
- b. Niemand kan aan de Westduitse bondskanselier de heen-
 nobody can to the West German president the to
 en terugreis voorschrijven.
 and from_journey prescribe
*Nobody can prescribe both ways of the journey to the West Ger-
 man chancellor.*

The non-canonical word order is by far the least frequent realization of the dative construction, with less than five percent of the data falling in this category. It is nevertheless possible to identify certain factors that increase the chance of finding this alternant. We first investigate the influence of weight on the order of the NP and PP argument.

3.5.1 Weight

If we simply count the average number of words for all shifted and canonical PP recipients, we find that the canonical PP recipients are much heavier than the shifted PPs, as shown in table 3.6. In addition, the canonical direct objects are lighter. This is in line with the principle 'light precedes heavy'.

However, we need to distinguish the relative ordering in the midfield from extraposition phenomena, which are known to be influenced by syntactic weight. In V2 sentences, it is impossible to see whether the second pole, the right edge of the midfield follows or precedes the PP. Any sentence is thus ambiguous between a canonical word order sentence and an instance of PP extraposition. To ensure that the PP is in the midfield and has not been extraposed, we only include sentences in which the PP is followed by

Dataset			OBJ1	OBJ2	OBJ1/OBJ2
CGN	NP PP	(N=56)	2.02	1.93	1.05
Alpino	NP PP	(N=17)	2.53	2.29	1.10
TwNC	NP PP	(N=100)	3.27	2.17	1.51
CGN	PP NP	(N=7)	4.71	2.71	1.73
Alpino	PP NP	(N=8)	3.50	3.25	1.08
TwNC	PP NP	(N=101)	2.64	2.70	0.98

Table 3.7: Average number of words of non-extrapolated PP recipients.

a verbal cluster, which indicates the right pole. Controlling for this factor, the weight effect disappears. However, the amount of data we have is very small, especially for the inverted word order. This is caused by the effect of combining various constraints on the data. PP recipients are less frequent than NP recipients and the shifted word order is less frequent than the canonical word order. If we combine these constraints and furthermore restrict ourselves to unambiguously non-extrapolated PPs, we are left with very little data. This data sparseness can be overcome with automatically annotated data. We extracted all shifted PP patterns from the automatically annotated TwNC corpus and checked the results manually. This resulted in a total of 101 examples. We furthermore extracted the first 100 correct examples of the unshifted construction and compared the average weight of direct and indirect object in these test sets. The results in table 3.7 show that direct object and indirect object are almost equally heavy. Comparing the shifted and the unshifted construction, we see that the direct objects are heavier in the shifted construction than in the canonical argument order, and the indirect objects are lighter in the shifted construction than in the canonical argument order. This is contrary to what is expected, based on the complexity principle.

The large difference between the weight of unambiguously non-extracted PP recipients and those that may have been extracted shows that the influence of weight on argument ordering is not a linearization constraint: it did not influence the ordering of the two objects in the middle field and it does not influence the relative order of the NP and the PP in the middle field in dative PP constructions. Instead, it is a preference for heavy constituents not to be embedded in the VP, but to be extrapolated instead.

Dataset	OBJ1			OBJ2		
	Indef	Def	Prons	Indef	Def	Prons
TwNC NP PP (N=100)	44%	56%	6	22%	78%	5
TwNC PP NP (N=101)	85%	15%	0	36%	64%	0

Table 3.8: Pronouns, definites and indefinites in the dative PP construction.

3.5.2 Pronominality and definiteness

That leaves open the question of what influences the relative order of the NP and the PP argument. The obvious candidates: information structure and pronominality. The first has often been proposed in the literature as a factor influencing word order (Gundel, 1988; Prince, 1992; Arnold et al., 2000) and the second (related) factor was already shown to be the crucial factor determining the relative order of the objects in the middle field of the double object construction. Table 3.8 list the numbers of pronouns, definites and indefinites in the shifted and unshifted (manually checked) datasets.

The table shows that the percentage of indefinite direct objects is much higher in the shifted OBL OBJ1 construction (85%) compared to the canonical OBJ1 OBL construction (44%), in line with the principle that says definite precedes indefinite. However, it would be incorrect to conclude that we have a constraint DEF<INDEF or NEW (Choi, 1996) that outranks the constraint on canonical wordorder. It is true that in a little over over half the non-canonical examples (55%), the direct object was indefinite and the indirect object was definite, in which case the non-canonical word order could be accounted for by the definiteness principle. But 32% of the canonical examples also had indefinite direct objects and definite indirect objects. And since the canonical construction is many times more frequent than the shifted construction, this means that the majority of the examples with an indefinite OBJ1 and a definite OBJ2 examples is in the canonical argument order. Instead we do assume a constraint NEW, but it is generally ranked below CANON. The distance between the two constraints must be small, though, so that there is some probability of NEW outranking CANON under a stochastic implementation of OT. This constraint then also explains why we find no direct object pronouns in the non-canonical construction: even if NEW outranks CANON, it simply does not apply to pronouns, as they are by definition old information.¹⁰

(16) NEW: [-NEW] should precede [+NEW] Choi (1996)

¹⁰It does not, however, account for the lack of pronominal *indirect* objects. We leave this issue for future research.

Among the non-canonical examples, we find an interesting subgroup. A number of examples consists of expressions containing semantically light verbs which form a collocation with their direct object: *aandacht geven/schenken*, *excuses aanbieden*, *gehoor geven*, *uitdrukking geven*, *leiding geven*, ‘give attention’, ‘offer excuses’, ‘give attention’, ‘give expression’, ‘give leadership’ (17) and this tendency of the direct object to appear at the right pole of the midfield appears to increase with the strength of the collocation. We can illustrate this by looking at two examples which mean almost the same, *aandacht geven/schenken* and *gehoor geven* ‘give attention’. *Gehoor* is in the lexicon as an independent noun, but it is most often used with the verb *geven* ‘to give’ and thus forms a very strong collocation with *geven*. This is not true for *aandacht*, which forms a weak collocation with the verbs *geven* and *schenken* ‘to give’. The stronger collocation appeared twice in canonical order in the complete 75M word corpus vs. 6 times in the shifted construction, the weaker collocation appeared 40 times in the canonical word order vs. 17 times shifted. This is as predicted by the Inherence Principle (Haeseryn et al., 1997) and the idea of semantic connectedness (Wasow, 2002) in section 3.2: the closer the verb and the direct object are connected, the higher the chance of finding the non-canonical order in which the direct object immediately precedes the verb.

- (17) Het is zeer uitzonderlijk dat het Israëlische leger in een dergelijk
 it is very exceptional that the Israeli army in a such
 geval aan nabestaanden zijn excuses aanbiedt.
 case to relatives his excuses offers
*It is very exceptional for the Israeli army to offer an apology to the
 relatives in such a case.*

To conclude, the dative PP construction is most frequently realized in the canonical word order. Two factors were shown to increase the chance of finding the non-canonical order: definiteness and light verb constructions. However, it is not generally the case that in those restricted contexts, the non-canonical variant is more frequent than the canonical one. Expressed in OT terms, we must conclude that the constraint on canonical word order is the highest ranked relevant constraint, but certain lower ranked constraints are close enough to outrank canonical word order with some (low) frequency. There remain some non-canonical examples in which neither of these factors is present. We leave the question as to why we find the inverted order in these cases for further research.

In the last two sections, we looked at the order of the arguments in the double object construction and in the recipient PP construction. It was

shown that pronominality determines the order in the double object construction and that definiteness and “inherence” influence the argument order in the dative PP construction. The next section deals with the question what determines the choice for an NP or a PP recipient.

3.6 The NP/PP Alternation

The third and last alternation in the Dutch dative construction concerns the choice for an NP or PP recipient argument. As the syntactic category does not have to influence the linearization of arguments in Dutch, we do not expect to find influence from general linearization principles here. Instead, we expect to find constraints that directly influence the realization of the recipient. One such constraint is the selectional restriction of verb classes, as proposed in Levin (1993).

3.6.1 Lexical preferences

Levin (1993) argues that some English verb classes select for one variant in the dative alternation and some verbs select for the other. As the argument order variation and the NP/PP alternation go together in English, these selectional restrictions influence both argument order and the syntactic function of the recipient. In Dutch, we would expect that such construction specific constraints only influence the realization of the recipient argument, not the order of the arguments, which is assumed to be governed by more general linearization constraints.

We tested the influence on verb class on the argument variation in the double object construction on the one hand and the NP/PP alternation on the other hand by calculating the association between verb lexeme and word order/recipient type in the annotated Alpino and CGN corpora. We express this association with a log-likelihood score (see section 1.2.3 for a description of the log-likelihood measure) and regarded each verb as a separate class. As we already saw that the argument order is influenced by the syntactic category of the direct object, we controlled for this by calculating the influence separately for *het*-objects, pronominal objects and NP objects. Table 3.9 summarizes the log-likelihood scores for both alternations in all three object classes. We see that the association between verb lexeme and word order is not significant in any of the OBJ1 classes. The association between verb lexeme and the NP/PP alternation does reach significance in two classes out of three classes, suggesting that the verb lexeme does influence the choice for an NP or PP recipient. These results tell us that the distribution of NP and

Alternation	Obj1	Degrees of Freedom	LL	Significant
Arg Order (NP NP)	NP	35	6.2	no
	pron	20	22.9	no
	het	7	4.4	no
NP/PP Alternation	NP	40	79.8	p=0.001
	pron	24	36.5	p=0.050
	het	7	8.3	no

Table 3.9: Loglinear (LL) scores for the relation between verb lexeme and surface form in different OBJ1 categories.

Alternation	Ent before	Ent Cat	Ent Verb	Ent both
NPNP order	0.172	0.110 -36%	0.152 -12%	0.094 -45%
NP/PP recipient	0.578	0.578 -0%	0.426 -26%	0.422 -27%

Table 3.10: Entropy of the word order and dative NP/PP alternation

PP recipients over the different verbs cannot be attributed to chance. This does not necessarily mean that it is due to the different lexical preferences.

To get a better idea of the relative impact of direct object category on the one hand and the verb lexeme on the other hand on the argument order and the NP/PP alternation, we calculated the entropy of the system, based on the automatically parsed TwNC corpus. The results are in figure 3.10. This entropy is a measure of the uncertainty about whether or not shift will apply (or whether to realize the recipient as an NP or a PP). We first calculated the entropy of the system without adding any information. This starting entropy is low, as the canonical word order is much more frequent than the shifted word order. We then added either information about the syntactic category of the direct object (*het*, personal pronoun or NP) or the verb lexeme. The entropy reduction after adding OBJ1 information is much higher than after adding verb lexeme information, even though there are many more verb lexeme categories than OBJ1 categories. We did the same calculations for the NP/PP alternation. Here, we see the reverse picture: adding OBJ1 category information does not reduce the entropy of the system, but adding verb lexeme information leads to an important entropy reduction. Adding information about both the verb lexeme and the direct object category leads to an entropy reduction as big as the sum of the entropy reductions of the two pieces of information individually. This indicates that the OBJ1 category and the verb lexeme are independent variables.

Adding verb class information reduced the entropy considerably, confirming our hypothesis that the construction specific lexical preferences only influence the NP/PP alternation, not argument ordering. The entropy did not go down to zero, but it is not expected to, as some verbs allow free variation. Nevertheless, it is not excluded that other factors influence the choice for one NP recipients over PP recipients or vice versa. We check the influence of two more general constraints: weight and pronominality.

3.6.2 Weight and pronominality

If we look at the average weight of indirect objects in table 3.5 on page 85, we see that the recipient arguments that are realized in PPs are much heavier than those that are realized as NPs. But these numbers may not be measuring an effect of weight on the NP/PP alternation proper: they may be a side effect of the constraints on DOS identified in section 3.4. We saw that even in Dutch, word order and the NP/PP alternation are not completely independent. In the double object construction, the theme-recipient order is available for pronominal themes only. Thus, in sentences with heavy recipients and full NP themes, realization as a PP is the only way in which the recipient can be “moved” to the right.

In order to look at the NP/PP alternation proper, we have to control for this and ensure that the order of the arguments in the sentence does not play a role. We decided to restrict ourselves to PP recipients which co-occur with an inanimate pronominal direct object. In these sentences, the direct object may precede both an NP and a PP indirect object in the midfield. We thus controlled for influences of ordering effects. Furthermore, we only included PP recipients which undeniably precede the second pole (i.e. precede the verb cluster). This excludes extraposition effects.

For the annotated data (Alpino and CGN), that left us with only 13 examples. To collect a more representative data collection, we extracted from the automatically parsed TwNC corpus 200 sentences that met our criteria, and manually checked them for parse errors. For NP recipients, we constructed a similar dataset. For the annotated data, we took the CGN data, as 12 of the 13 PP examples were from CGN. In addition, we again extracted 200 examples from the automatically annotated corpus, again restricting ourselves to double object constructions with pronominal objects. The table in 3.11 still shows a striking difference in weight between NP and PP recipients

Although we know now that the difference in weight must be associated with the difference in syntactic category, not with the order of the arguments it still does not prove that there is an influence of the Complexity Principle on the NP/PP alternation. The distributional differences in the double object

Dataset	NP		PP	
Alpino+CGN	1.12	N=52	1.53	N=13
TwNC	1.43	N=200	2.35	N=200

Table 3.11: Average weight for PP and NP recipients in the midfield

construction appeared to be weight-driven, but turned out to be based on pronominality. The same applies to the dative PP construction. There were 20 pronouns among the PP recipients vs. 161 among the NP recipients. If we remove pronominal recipients from the data (leaving 180 sentences in the PP category and only 39 in the NP category) we find that the PP recipients are in fact lighter (2.5 words) than the NP recipients (3.1 words). Weighing of the non-pronominal NP recipients on a larger set of 100 instances reduced the average to 2.7 words, which is however still heavier than the PP recipients. The conclusion must be that there is not so much a difference in weight between NP and PP recipients, but rather a pronominality difference. We did one more test to illustrate the effect of pronominality on the NP/PP alternation: we compared the percentage of one word recipients taken up by pronouns in the automatically parsed corpus (thus controlling for weight effects). We find 22% (20041/92553) pronouns in the one word NP recipients vs. 6% (347/5527) in the PP data. In the annotated CGN data this contrast is less pronounced, but still clearly there: 82% (581/706) vs. 49% (52/106). All together, this shows that pronominal recipients disprefer realization as an oblique. This effect cannot be reduced to the general linearization constraints we saw in section 3.4 (PRON<NP): in our data, all direct objects were pronominal, so that the order with the recipient following the theme was possible both with an NP and a PP recipient. Many of our examples were in fact shifted. In other words: in this restricted domain, realization as an NP or a PP does not determine the relative order of the arguments.

Finally, recall from the literature section that Bresnan and Nikitina (2003) argued that local person NPs should be realized as objects, not obliques. As a local person NP is always realized as a pronoun, this could explain why we find many more pronominal NP recipients than pronominal PP recipients. But no evidence for such a restriction was found: of the 20 pronouns in our 200 sentence PP test set, 12 (60%) were local. Of the 161 pronouns in the NP test set, 84 (52%) were local.¹¹ We also tested the existence of

¹¹Note that third person inanimate PP recipients were excluded. These are realized as R-pronouns, which we filtered out (see section 3.3.1). This could potentially influence the data, especially since Bresnan et al. (2005) argue that inanimate recipients prefer the PP structure. However, as the vast majority of recipients is animate (in fact, *all* NP

	local	3rd person
NP NP _{unshifted}	101	52
NP NP _{shifted}	13	9
NP PP	7	3
PP NP	2	1

Table 3.12: Person features of pronominal indirect objects

a person effect in our annotated data (without any further restrictions on direct object category or extraction). Table 3.12 shows the distribution of third person and local pronouns over the four alternants. The differences between the frequencies of third and local person pronouns is not significant ($p=0.05$). Even after aggregating together the counts for the two orderings in the double object construction and the two orders in the dative PP construction, we still did not reach significance. Although the numbers are too small for definite conclusions, they do suggest that person does not influence the NP/PP alternation.

3.6.3 Implementation in OT

We have identified two factors that influence the dative NP/PP alternation in Dutch: lexical preferences and pronominality. In addition, we saw in earlier that the dative PP construction is less frequent, over all. In this section, we incorporate these findings in our OT model. First, we look at Bresnan and Nikitina’s account of the English alternation (Bresnan and Nikitina, 2003). They adopt the set of constraints in (18).

(18) Constraints on the Dative Alternation (Bresnan and Nikitina, 2003)

*STRUCT: avoid syntactic structure (here: *PP)

FAITH(REC): express the recipient role of a verb with distinct marking (case or adposition)

OO-PRIMACY: OBJ2 strictly dominates OBJ1 on hierarchies of informational prominence.

recipients were animate), we do not expect that ignoring this factor has a large influence on the person effect.

HARMONY(1,2): *NP_{Noun} & *PP_{1,2Person}.

*STRUCT can be applied to both English and Dutch to account for the skewed overall distribution of NP and PP recipients; we simply adopt this constraint. We have found no evidence of an influence of OO-PRIMACY on the NP/PP alternation in Dutch. However, pronominality and information structure did play a role in the linearization of arguments. It would be interesting to see how many of the effects that have been analyzed as OO-PRIMACY effects could be accounted for with our linearization constraint, given the strict argument ordering in English. However, if the constraint proves indispensable, this is fully compatible with our account for Dutch (which would have the constraint ranked below all relevant constraints).

This is not the case for HARMONY(1,2). Bresnan and Nikitina's harmony constraint penalizes pronominal recipients realized as a dative PP, which is what we want for Dutch, too. However, it also penalizes realizations for which we found no evidence of dispreference in Dutch, such as full NP OBJ2s. We propose therefore to have the simplex constraint *PP_{pro} instead.¹² Note that our account excludes the existence of the harmony constraint. By definition, constraint conjunctions outrank both their conjuncts, but our data contradict a constraint HARMONY(1,2) which outranks *PP: such a constraint would predict a dispreference for full NP indirect objects, which we did not find. Our account thus clashes with the account of Bresnan and Nikitina under the assumption of a universal set of constraints. It is an interesting question whether the general linearization constraints we identified can be used to account for the English data without the use of HARMONY(1,2). We leave this for future research.

That leaves the question of how to implement the lexical preferences. Bresnan and Nikitina assume that FAITH(REC) is parameterized for different classes of ditransitive verbs. These parameterized constraints are then ranked at various positions in the hierarchy, as in the constraint ranking in (19) (incorporating only a few of the lexical preferences), so that different frequencies are predicted for the PP and NP realizations of their recipient arguments.

- (19) Constraint Ranking for the Dative Alternation (Bresnan and Nikitina, 2003)

¹²This constraint could be derived from aligning the pronominality hierarchy with the Core and Noncore argument hierarchy in the same way the two conjuncts of the HARMONY(1,2) constraint were derived, using the techniques familiar from the work of Aissen (1999, 2003).

$$\begin{aligned} & \text{OO-PRIMACY} \gg \text{FAITH}(\text{Rec})_{yell} \gg \text{HARMONY}(1,2) \\ & \gg \text{FAITH}(\text{Rec})_{fax}, \text{FAITH}(\text{Rec})_{give} \gg * \text{STRUCT} \end{aligned}$$

We adopt this approach in order to illustrate how lexical preferences interact with the other constraints. However, some remarks should be made about this approach. The parameterized faithfulness constraints are unlikely to be universal. As such, they are in violation of the basic assumption in OT that the set of constraints is universal and that the only source of variation is the ranking of these constraints (Prince and Smolensky, 1993). At the same time, Smolensky and Legendre (2005) acknowledge that this principle may have to be weakened to account for certain language particularities. An alternative is to introduce a lexical feature which specifies the strength of the recipient. The constraint could then refer to this feature instead of the verb (class) itself. Such an approach crucially relies on a lexicon friendly OT model (van der Beek and Bouma, 2004). The two approaches make different predictions: in the first model, all members of a (semantically motivated) class have the same distribution, whereas the latter allows variation within such classes. There appear to be such differences, but at the same time, Lapata (1999) showed that Levin's verb classes do have empirical value. This regularity would be unaccounted for in the second model.

In tableau 3.13, we illustrate how lexical preferences interact with the other constraints. We included only one constraint FAITH(REC), which favors the PP realization. We rank it on a par with *PP, so that free variation is expected if no other constraint penalizes one of the constructions. An example of such a verb is *betalen* 'pay'. We expect a 50-50 division between NP and PP recipients if both arguments are NPs. If we have a pronominal recipient, *PP-PRO prevents it from being realized as a PP resulting in a double object construction.

- (20)
- a. Bo betaalt de student tien euro.
Bo pays the student ten euros
 - b. *Bo betaalt tien euro de student.
Bo pays ten euros the student
 - c. Bo betaalt tien euro aan de student.
Bo pays ten euros to the student
 - d. ??Bo betaalt aan de student tien euro.
Bo pays to the student ten euros
 - e. ??Bo betaalt de student het.
Bo pays the student it

- f. ??Bo betaalt het de student.
Bo pays it the student
- g. Bo betaalt het aan de student.
Bo pays it to the student
- h. Bo betaalt aan de student het.
*Bo pays to the student it
- i. Bo betaalt hem tien euro.
Bo pays him ten euros
- j. *Bo betaalt tien euro hem.
Bo pays ten euros him
- k. ??Bo betaalt tien euro aan hem.
Bo pays ten euros to him
- l. ??Bo betaalt hem tien euro.
Bo pays to him ten euros
- m. ??Be betaalt hem het.
Bo pays him it
- n. Bo betaalt het hem.
Bo pays it him
- o. ??Bo betaalt het aan hem.
Bo pays it to him
- p. *Bo betaalt aan hem het.
Bo pays to him it

At least two other groups of verbs exist. Verbs like *verwijten* ‘blame’ prefer the double object construction, unless there is some external reason for realizing the recipient as a PP, for example to avoid ambiguity (see section 3.7). For these verbs, the faithfulness constraint either does not apply or it is ranked below the markedness constraint *PP. Finally, there are verbs such as *verkopen* ‘sell’, which generally prefer to realize the recipient argument as a PP, unless it is a recipient. Here, the appropriate FAITH(REC) must outrank the markedness constraint *PP. The minimal constraint ranking accounting for these three classes is thus as in (21).

(21) Minimal Constraint Ranking for the Dutch Dative Alterations

$$\begin{aligned}
 &(\text{PRO}_{it} < \text{NP}/\text{PRO}) \gg (\text{PRO}_{3rd} < \text{NP}) \gg *PP\text{-PRO} \gg \\
 &\text{FAITH}(\text{Rec})_{\text{verkopen}} \gg *PP, \text{FAITH}(\text{Rec})_{\text{betalen}} \gg \\
 &(\text{OBJ2} < \text{OBJ1} < \text{OBL}) \gg \text{NEW}
 \end{aligned}$$

Input: <i>pays</i> (<SUBJ><OBJ1><OBJ2>)		PRO _{it} <NP/PRO	PRO _{3rd} <NP	*PP-PRO	*PP	FAITH(Rec)	OBJ2<OBJ1<OBL	
OBJ1='ten euros' OBJ2='the student'	☞ NP NP _{unshift}					*		(20-a)
	☞ NP NP _{shift}					*	*!	(20-b)
	☞ NP PP				*			(20-c)
	☞ PP NP				*		*!	(20-d)
OBJ1='it' OBJ2='the student'	☞ NP NP _{unshift}	*!				*		(20-e)
	☞ NP NP _{shift}					*	*!	(20-f)
	☞ NP PP				*			(20-g)
	☞ PP NP				*		*!	(20-h)
OBJ1='ten euros' OBJ2='him'	☞ NP NP _{unshift}					*		(20-i)
	☞ NP NP _{shift}					*	*!	(20-j)
	☞ NP PP			*!	*			(20-k)
	☞ PP NP			*!	*		*	(20-l)
OBJ1='it' OBJ2='him'	☞ NP NP _{unshift}	*!				*		(20-m)
	☞ NP NP _{shift}					*	*	(20-n)
	☞ NP PP			*!	*			(20-o)
	☞ PP NP			*!	*		*	(20-p)

Table 3.13: Competition between the four alternants of the Dutch dative alternation.

3.7 More Factors in the Dative Construction?

We have focused on three factors that influence the dative alternations: weight, pronominality/definiteness and lexical preferences. No doubt many more factors influence the realization of ditransitive verbs in one way or another.

Bresnan et al. (2005) argue that animacy is a relevant feature. Unfortunately, none of the available corpora of Dutch is annotated with information about animacy, making it impossible to test this hypothesis on corpus data. Within the restricted search space of the pronominal recipients, there were too few inanimate recipients to draw any conclusions. That being said, it does seem to be the case that with (marked) inanimate recipients, the DOS is less acceptable and the PP-construction is preferred. This is illustrated in the constructed examples in (22).

- (22) a. Ik geef dit boek een tien.
 I give this book a ten
I give this book ten out of ten.
- b. ?Ik geeft dat geen enkel boek
 I give that no single book
I do not give that to any book.
- c. En toch geef ik dat wel aan dit boek.
 and still give I that indeed to this book
But I still do give that to this book.

The markedness of (22-b) may also be explained by the aim to avoid ambiguity. If both objects are inanimate, it is harder to tell which one is the direct object and which one the indirect. In such cases, the canonical word order or dative marking with a PP appear obligatory. Similarly, if the direct object is an atypical object, e.g. a local (1st or 2nd person) pronoun, it is easily misunderstood as an indirect object (which is often local). By marking the real indirect object with the preposition *aan*, this reading is excluded and ambiguity is reduced.¹³ This would explain the ungrammaticality of the constructed example in (23-a) and the grammaticality of (23-b).¹⁴

¹³This could also be regarded a mild OO-PRIMACY effect.

¹⁴Note that the shifted version (i) is also ungrammatical under the intended reading. This was already accounted for in section 3.4. Even if it were grammatical, it would not solve the ambiguity.

- (i) %als ik jou Ajax verkoop.
 if I you Ajax sell

- (23) a. *als ik Ajax jou verkoop.
 if I Ajax you sell
 b. als ik jou aan Ajax verkoop.
 if I you to Ajax sell
 If I sell you to Ajax.

There are indications that the surface string also has some influence on the realization of the recipient argument. Among the sentences with PP recipients, for example, we find many that have proper name recipients, proper name agents and non-pronominal themes. As DOS is only available for pronouns, only the canonical double object construction is possible. This canonical argument order would lead to two proper names in a row (24-a). Realizing the recipient as a PP argument successfully avoids this sequence of proper names (24-b).

- (24) a. Daar gaf volgens de overlevering God Mozes het
 there gave following the tradition God Moses the
 gebod "Gij zult niet stelen".
 commandment thou shalt not steal
 b. Daar gaf volgens de overlevering God aan Mozes het
 there gave following the tradition God to Moses the
 gebod "Gij zult niet stelen".
 commandment thou shalt not steal
 *Tradition has it that this is the place where God gave Moses the
 commandment "Thou shalt not steal".*

Finally, we argued in section 3.2 that the account of Reinhart (1996) in terms of stress and focus is based on dubious data and cannot account for all data. At the same time, we noted that there are reasons to expect an influence of focus on linearization. This influence is not restricted to the dative alternation and should be studied with a corpus that is annotated with information structure. Unfortunately, such a corpus is not available for Dutch.

In the previous sections, pronominality and definiteness constraints were shown to override the canonical word order in the Dutch dative alternation in some instances and the NP/PP alternation was shown subject to lexical preferences and—surprisingly—pronominality constraints. This section discussed some additional influences which appear to influence the dative alternation. The most important constraints were modeled in the OT framework. The next section provides some additional evidence for an important part of this model, namely the ordering constraints on nominals. This evid-

ence comes from a second construction which is sensitive to the pronoun scale: the Accusativus cum Infinitivo.

3.8 Additional Evidence: the AcI

In section 3.4 it was shown that a pronoun hierarchy exist in Dutch, similar to the one for German (Müller, 2001). It was furthermore shown that DOS is sensitive to this scale: the weaker the pronoun, the more prone it is to shift. This was formalized in a set of linearization constraints. In this section we will digress from the dative alternation to illustrate that the various constraints for aligning pronouns of different strengths can also be applied to other word order alternations. We show how these constraints account for the distribution of embedded object shift (EOS) in the Accusativus cum Infinitivo (AcI) construction.

The AcI construction illustrated in examples (25) and (26) in and figure 3.1 is headed by a sensory verb, the verb *laten* (to let) or the verb *helpen* (to help). The verb takes an object and an XCOMP. The embedded subject is functionally controlled by the object.

- (25) Op haar elfde zag ze Russische tanks haar land
 On her eleventh saw she Russian tanks her country
 binnenvallen.
 invade
At age eleven she saw Russian tanks invade her country.

Several LFG analyses of this construction exist, e.g. Bresnan et al. (1982), Zaenen and Kaplan (1995) and Kaplan and Zaenen (2003). All nominal arguments (also the embedded ones) are selected for in the VP, all verbal arguments in V', thus accounting for the crossing dependencies that occur when one AcI constructions is embedded in another, as illustrated in a well known example from the literature (26):

- (26) omdat ik Cecilia Henk de nijlpaarden zag helpen voeren.
 because I Cecilia Henk the hippos saw help feed
because I saw Cecilia help Henk feed the hippos.
- (27) C-structure rules for the AcI-construction Kaplan and Zaenen (2003)
- $$VP \rightarrow \left(\begin{array}{c} NP^* \\ \uparrow XCOMP^* \text{ OBJ} = \downarrow \end{array} \right) V'$$

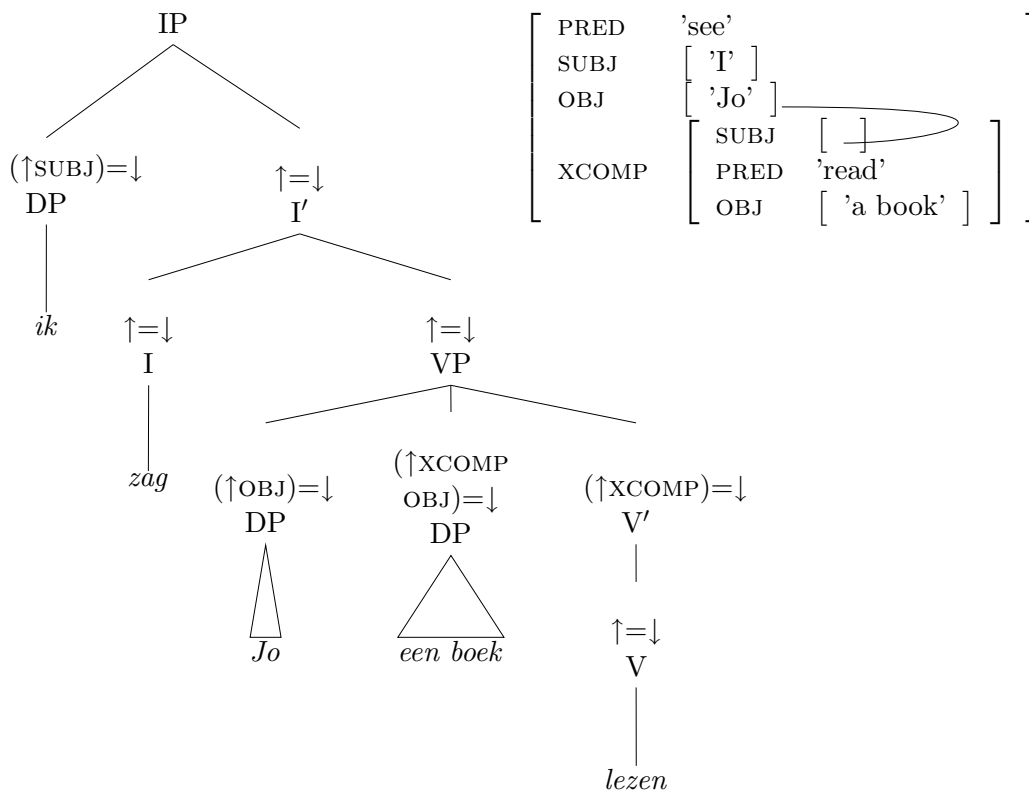


Figure 3.1: C-structure and f-structure for an AcI-construction in Dutch

$$V' \rightarrow V \left(\begin{array}{c} V' \\ (\uparrow \text{XCOMP}) = \downarrow \\ (\uparrow \text{XCOMP}^+ \text{OBJ}) \not\prec_f (\uparrow \text{OBJ}) \end{array} \right)$$

The order of the nominal arguments is restricted to the canonical word order in (25) and (26) by the f-precedence requirement $(\uparrow \text{XCOMP}^+ \text{OBJ}) \not\prec_f (\uparrow \text{OBJ})$ (Kaplan and Zaenen, 2003). This constraint states that the constituent that maps onto the embedded OBJ1 in the f-structure cannot precede the constituent that maps onto the f-structure of the main clause direct object. However, under certain conditions, the embedded object can shift over the higher object (or embedded subject) (28). In other words: the f-precedence constraint is violable. The conditions under which we find EOS resemble the conditions on DOS. A difference is that DOS was only blocked with local person pronouns, while EOS is blocked with all animate pronouns. This is best illustrated with animate and inanimate examples of the weak pronoun *ze* (them) (29-b)-(29-a). Note that inanimate objects are very unmarked. More marked objects have to stay in their canonical object position.

- (28) a. Ik zag 't Jo doen.
 I saw it Jo do
I saw Jo doing it.
- b. Ik zag 't 'r doen
 I saw it her do
I saw her doing it.
- c. Ik zag dat haar ouders doen.
 I saw that her parents do
I saw her parents doing that.
- d. Ik zag haar ouders dat doen.
 I saw her parents that do
I saw her parents doing that.
- e. Ik zag ze dat doen.
 I saw them that do
I saw them doing that.
- f. Ik zag het ze doen.
 I saw it them do
I saw them doing it.
- (29) a. Ik heb ze Jo door zien slikken.
 I have them Jo through seen swallow
I saw Jo swallowing them.
- b. %I heb ze Jo zien zoenen.
 I have them Jo seen kiss
I saw Jo kissing them.

The constraints on EOS resemble those on DOS. Again we see an interaction between pronouns that want to precede full NPs and canonical word order preventing this. In contrast to DOS, we now have the dividing line between animate and inanimate pronouns. This is modeled by ranking the canonical word order constraint for the AcI construction between the linearization constraints for animate and inanimate pronouns. The canonical word order constraint (30) states that complement functions (here restricted to OBJ1, OBJ2 and OBL) precede complement functions that are embedded in an XCOMP.

- (30) CF < XCF: OBJ1, OBJ2 and OBL precede (XCOMP OBJ1), (XCOMP OBJ2) and (XCOMP OBL)

The rest of the analysis works as for the OS in the double object construction, as illustrated in table 3.14: if both the OBJ1 and (XCOMP OBJ1) are full

Input: <i>see</i> (<SUBJ><OBJ1><XCOMP>)			PRO _{it} <NP/PRO	PRO _{3rd/inanim} <NP	CF <XCF	PRO _{3rd/anim} <NP	O2 < O1 < OBL	PRO _{loc} < NP
OBJ1='Jo'		☞ OBJ1 XOBJ1						
XOBJ1='a book'	ex.(25)	XOBJ1 OBJ1			*!			
OBJ1='Jo'		☞ OBJ1 XOBJ1	*!					
XOBJ1='it'	ex.(28-a)	XOBJ1 OBJ1			*			
OBJ1='her parents'		☞ OBJ1 XOBJ1		*!				
XOBJ1='that'	ex.(28-c)	XOBJ1 OBJ1			*			
OBJ1='them'		☞ OBJ1 XOBJ1						
XOBJ1='that'	ex.(28-e)	XOBJ1 OBJ1			*!			
OBJ1='them'		☞ OBJ1 XOBJ1	*!					
XOBJ1='it'	ex.(28-f)	XOBJ1 OBJ1			*			
OBJ1='Jo'		☞ OBJ1 XOBJ1				*		
XOBJ1='them'	ex.(29-b)	XOBJ1 OBJ1			*!			

Table 3.14: Embedded Object Shift in the AcI

NPs, we simply get the canonical word order, as none of the linearization constraints on pronouns fires. The same happens if both arguments are pronouns, but not local or *het*. However, if the direct object is *het* or it is another pronoun and the indirect object is a full NP, the linearization constraints on pronouns, which outrank the constraint on canonical word order, determine that we get an EOS.

Now what happens if we embed a ditransitive in an AcI construction? In principle both types of shift are possible: within the XCOMP argument, direct objects may shift in front of the indirect objects (DOS), and in addition, the embedded object may shift in front of the direct object of the main verb (EOS). Some of the possibilities are listed in (31). For three examples, we put the constraint interaction in tableau 3.15.

- (31) a. Ik heb de docent de leerlingen het boek zien geven.
 I have the teacher the students the book see give
I saw the teacher giving the students the book.




Input: <i>see</i> (<SUBJ><OBJ1><XCOMP>)		PRO _{it} <NP/PRO	PRO _{3rd/inanim} <NP	CF <XCF	PRO _{3rd/inim} <NP	O2 <O1 <OBL	PRO _{loc} <NP
OBJ1=NP XOBJ1=NP XOBJ2=NP (31-a)	 OBJ1 XOBJ2 XOBJ1 OBJ1 XOBJ1 XOBJ2 XOBJ2 OBJ1 XOBJ1 XOBJ2 XOBJ1 OBJ1 XOBJ1 XOBJ2 OBJ1 XOBJ1 OBJ1 XOBJ2			*! *!* *!* *!		*! * *	
OBJ1=NP XOBJ1=het XOBJ2=NP (31-b)	 OBJ1 XOBJ2 XOBJ1 OBJ1 XOBJ1 XOBJ2 XOBJ2 OBJ1 XOBJ1 XOBJ2 XOBJ1 OBJ1 XOBJ1 XOBJ2 OBJ1 XOBJ1 OBJ1 XOBJ2	*! *! *! *!		* ** **! *		* * *	
OBJ1=NP XOBJ1=NP XOBJ2=pron (31-c)	 OBJ1 XOBJ2 XOBJ1 OBJ1 XOBJ1 XOBJ2 XOBJ2 OBJ1 XOBJ1 XOBJ2 XOBJ1 OBJ1 XOBJ1 XOBJ2 OBJ1 XOBJ1 OBJ1 XOBJ2			*! *!* *!* *!	* *	*! * *	

Table 3.15: Tableaux for (31)

- b. Ik heb het de docent de leerlingen zien geven.
 I have it the teacher the students see give
I saw the teacher giving it to the students.
- c. Ik heb de docent hun het boek zien geven.
 I have the teacher them the book see give
I saw the teacher giving them the book
- d. %Ik heb hun de docent het boek zien geven.
 I have them the teacher the book see give

The model predicts that it is possible to shift an inanimate (XCOMP OBJ2) in front of the direct object of the main verb. Example (32) shows that this

results in bad sentences, contrary to what we expect. The ungrammaticality of (32) may be explained by the fact that inanimate indirect objects are rare. One may expect a ‘worst of the worst effect’ (Smolensky, 1995; Lee, 2003), resulting in the ungrammaticality of the combination of both a marked indirect object and a marked argument order.¹⁵

- (32) a. Ik zie hem het boek een tien geven
 I see him the book an A give
I see him give the book an A.
- b. ??Ik zie het hem een tien geven
 I see him that an A give
I see him give it an A.

3.9 Conclusion

We investigated the influence of various factors on the dative alternation in Dutch, all of which have been claimed to influence the dative alternation in English. These factors are weight, definiteness, pronominality and lexical preferences. The first three factors are considered general linearization principles. These linearization principles may influence the dative alternation in English, as the order of the arguments alternates with the syntactic category of the recipient. This is not the case in Dutch. The Dutch data thus allowed us to study the argument order and the NP/PP alternation in isolation.

We expected to find a split in the factors influencing the different aspects of the dative alternation in Dutch: linearization constraints influencing the argument order alternations, and construction specific constraints influencing the NP/PP alternation. This expectation is partially borne out. Lexical preferences of the verb indeed only influenced the choice for an NP or a PP recipient, not the order of the arguments. And pronominality and definiteness were indeed shown to influence the order of the arguments.

But not all expectations were borne out. Pronominality was assumed to be a linearization constraint, related to the definiteness constraint. It was thus predicted to have an influence on argument order but not on the NP/PP alternation. Corpus data however showed that pronominality does influence the NP/PP alternation (as well as argument order). This is in line with work based on harmonic alignment of the nominal hierarchy and the semantic role hierarchy (Silverstein, 1976; Aissen, 1999; Bresnan et al., 2005, for example).

¹⁵In an OT model, such effects may be modeled by means of a constraint conjunction. In this case, a conjunction of a (low ranked) constraint penalizing inanimate objects and a constraint penalizing non-canonical word order. The conjunction of two constraint always outranks both component constraints.

Syntactic weight is another classic linearization constraint. But surprisingly, syntactic weight did not influence the order of the arguments in the midfield, neither in the double object construction nor in the dative PP construction. It was shown that extra weight does increase the chance on finding extraposition.

Pronominality and definiteness did influence the argument order in the dative alternation, as expected. Pronominality and pronoun type determined whether DOS applied and definiteness was shown to have a mild influence on the relative order of the direct object and the dative PP. In general, there was a strong preference for the canonical argument orders, independent of the syntactic category of the recipient.

Formalizing the constraints on argument order in an OT setting allowed us to illustrate the ranking and interaction of the constraints. We thus showed how the most frequent patterns could be predicted. But the dative alternation shows a lot of variation, with different realizations occurring with extremely skewed distributions. Classic OT cannot account for these patterns. We sketched how a stochastic interpretation of OT, STOT Bowersma and Hayes (2001), could account for those less frequent realizations of the dative construction. It would be interesting to see whether the frequencies predicted by an actual implementation in STOT would match the frequency distributions that we observed in the corpora.

The NP/PP alternation was shown to be subject to lexical preferences. It remains an open question how these preferences are best modeled in OT. Lexical variation appears to be a serious problem for the general assumption in OT that all variation is driven by the ranking of universal constraints.

