

Optional Scrambling and Interpretation*

Introduction

It has often been observed in the literature that when a language allows for scrambling (which I will use here as a descriptive term for the occurrence of an object to the left side of an adverb), definite and other strong NPs freely scramble, whereas indefinite and other weak NPs are subject to certain restrictions. In De Hoop (1992) I proposed that only NPs of a certain semantic type (the generalized quantifier type) can scramble. Strong NPs such as *alle krakers* 'two squatters' in (1) are always of this type and therefore they scramble freely. Weak NPs can be lifted to the type of a generalized quantifier, but then they get a strong reading.¹ Thus, a weak NP such as *twee krakers* 'two squatters' will get a strong reading in (1b) (either partitive 'two of the' or referential 'those two', which indicates that its type is lifted to a generalized quantifier type, whereas in (1a) it can have either a weak (existential, predicative) or a strong (partitive or referential) reading.

* The research for this paper was supported by the Netherlands Organization for Scientific Research, NWO (grant 300-75-020). I thank the audience at the KNAW Colloquium *Interface Strategies*, September 1997, for helpful discussion. Two reviewers of this volume kindly commented upon a previous draft. Peter Ackema opened my eyes to certain problematic aspects of my original analysis. All remaining errors are my sole responsibility.

¹ In De Hoop (1992) it is argued that certain syntactic contexts trigger strong readings on weak NPs, where strong can be referential, partitive, generic, or generic collective, dependent on the nature of the determiner, the type of predicate and other syntactic factors, such as the choice of the adverb in (1). Clearly, the reason I call these readings *strong* is because they pattern with strong NPs in syntactic behaviour. Accordingly, strong NPs scramble freely in Dutch.

- (1) a. dat ik gisteren alle krakers / twee krakers heb gesproken
 that I yesterday all squatters / two squatters have spoken
 b. dat ik alle krakers / twee krakers gisteren heb gesproken
 that I alle krakers / two squatters yesterday have spoken
 “that I talked to all squatters / two squatters yesterday”

Note that I did not propose in my dissertation that scrambling is actually triggered by anything. That is, both strong NPs and weak NPs on a strong reading do not scramble obligatorily: in (1a) we find strong NPs as well as strong readings in the unscrambled object position. But when NPs scramble, they must be of a certain type. That idea is still maintained and elaborated upon in Van der Does and De Hoop (1998). In Van der Does and De Hoop (1998) we account for the fact that certain indefinites such as the one in (2), for example, are not allowed to scramble. This is so because they cannot be lifted to the type of a generalized quantifier in this type of context (the context of a light, i.e., a non-contrastive verb):

- (2) omdat iedereen (*een plas) nog *(een plas) moet doen
 since everyone still a piss must do
 “since everyone still has to take a piss”

The precise analysis is not important here. What is important is that, obviously, not all NPs can actually scramble. An indefinite such as *een plas* ‘a piss’ in *een plas doen* ‘take a piss’ can only be in unscrambled position. It is also true that weak NPs that do scramble necessarily get particular readings, readings that have been attributed in the literature to characteristics such as generalized quantifierhood, topicality, anaphoricity, or prominence. That is, in scrambled position, a weak NP usually gets a strong reading. In fact, it can only get a weak (existential) interpretation, if it is contrastively focused (cf. Choi 1996).

In this paper I will not be concerned with the conditions under which weak NPs such as indefinites can scramble. Instead, I will focus upon a related question that got a lot of attention in recent analyses of scrambling, and that is whether NPs that do scramble share a certain feature. That is, are there any features of either the object or any other element in the sentence or even the context, that actually *trigger* scrambling? I will argue that there is no independently motivated feature that does. That is, the strong NP in (1) optionally scrambles and so do other NPs that may scramble, in particular definites. Many have pointed out that scrambling is related to the structure of the surrounding discourse such that anaphoric NPs tend to scramble, whereas non-anaphoric NPs tend to stay in situ. I will not deny the existence of these correlations. What I will deny is the claim that all definites in scrambled position are indeed anaphoric as well as the claim that definites in unscrambled position cannot be anaphoric. In fact, these claims seem easy to falsify. Consider (3) below as an example of a referential, anaphoric, topical/presuppositional, D-linked, familiar definite that freely scrambles:

- (3) a. Heb je Jane (het geld) gisteren (het geld) gegeven?
 have you Jane (the money) yesterday (the money) given
 “Did you give Jane the money yesterday?”

- b. Ja, ik heb Jane (het geld) al eergisteren (het geld)
 yes I have Jane already (the money) the-day-before-yesterday (the money)
 gegeven
 given
 “Yes, I gave Jane the money yesterday already”

On the one hand, the definite that is introduced in the question and that is obviously functioning as an anaphor in the answer, does not have to scramble in (3b). On the other hand, when it is introduced in this out of the blue question, it can be focused and nevertheless scramble and give a well-formed result in (3a). Scrambling the definite in (3) is truly optional in both the question and the answer. Moreover, non-referential, non-anaphorically distressed, non-contrastively focused, dependent definites that combine with a light (non-contrastive) verb, freely scramble as well:

- (4) omdat iedereen (de was) nog (de was) moet doen
 since everyone (the laundry) still (the laundry) must do
 “since everyone still has to do the laundry”

For a semantic account of the difference in scrambling possibilities between definites and indefinites in contexts like these (with light or non-contrastive verbs), I refer the interested reader to Van der Does and De Hoop (1998). At this point, simply observe that the definite in (4) optionally scrambles, irrespective of its discourse status (it does not have to be anaphoric). The indefinite in (2), on the other hand, is not allowed to scramble, again irrespective of its discourse status (we cannot reach the well-formedness of the scrambled variant of (2) by anaphoric distressing of the indefinite object). The fact that non-anaphoric definites optionally scramble (cf. (4)), just like other definites (cf. (3)), and unlike predicative indefinites (cf. (2)), indicates that a proper analysis of scrambling should not be based on discourse features reflecting topic-focus structure or prominence, nor on phonological features reflecting anaphoric distressing, nor on semantic features reflecting referentiality, specificity, or presuppositionality. Moreover, in languages with a difference in two types of structural case for objects, it is not the case that elements bearing strong case obligatorily scramble either (cf. De Hoop 1992, Butt and King 1996, a.o.). I conclude that scrambling is not *driven* at all. Scrambling is in principle optional. When certain elements do not optionally scramble, this might be due to other, independent principles. Indefinite objects of light verbs, for example, have to be semantically incorporated (cf. Van Geenhoven 1996, Van der Does and De Hoop 1998). Similarly, the fact that clitics and weak pronouns in general have to scramble in languages like Dutch, may be explained by another, independent (presumably syntactic) restriction as well.

1. Word order variation and context

Recent approaches to scrambling phenomena argue that in apparent cases of optional scrambling, there is in fact no (true) optionality. Diesing and Jelinek (1995) argue that

referential definite NPs obligatorily scramble in order to escape a mechanism of existential closure that is applied to the VP. They claim this requirement to follow from a more general condition that requires the relative scope of operators to be syntactically fixed. However, their examples with definites in non-scrambled position are in fact well-formed:

- (5) weil ich selten die Katze streichle
since I seldom the cat pet
“since I seldom pet the cat”

The theory of Diesing and Jelinek predicts the sentence in (5) to be ungrammatical. But, as a matter of fact, it is not ill-formed at all, which also holds for the Dutch translation of this sentence. Nevertheless, Diesing and Jelinek mark (5) with the grammaticality indication *?, in their own words “*to indicate markedness in the sense that some contrastive context is required for felicity*”. Evidently, then, definites do not obligatorily scramble. The next question is whether indeed some special context is required to allow for a referential definite in unscrambled position or alternatively, for a non-topical, non-anaphoric definite in scrambled position. I will show that the answer to this question is simply ‘no’: context does not impose any restrictions on which word order variant may be used in the case of definites. My analysis crucially differs from Neeleman and Reinhart (1998) and Choi (1996) in this respect. These authors recognize that scrambling is optional in many cases, but argue that there can be no true optionality in the sense that word order variants differ in how optimal they are in a certain context.

At this point, consider the Dutch pair of sentences in (6):

- (6) a. omdat ik zelden de kat aai
since I seldom the cat pet
b. omdat ik de kat zelden aai
since I the cat seldom pet
“since I seldom pet the cat”

The two sentences in (6) are both well-formed. The question is whether we can find a difference in interpretation between the two. Williams (1997) argues that when a language shows scrambling, then there must be a difference in meaning between the scrambled and the unscrambled form, due to a generalized *Blocking Principle*. So, is there a difference in interpretation between (6a) and (6b)? At first sight, there does indeed seem to be a difference: in (6b) the cat is readily interpreted as an anaphor (the cat is already present in the discourse), whereas in (6a) the most unmarked interpretation involves a non-anaphoric cat (the cat uniquely refers, for instance, to the speaker’s cat, but it is new in the discourse). However, note that these different interpretations for (6a) and (6b) arise in the absence of any other phonological or contextual clues.

Accordingly, Neeleman and Reinhart (1998) and Choi (1996) would both predict that the distribution of the word order variants in (6) is not absolutely free: the use of one variant will be obligatory in certain actual contexts, and impossible in certain others. If the examples in (6) are embedded in actual contexts, however, that prediction is not borne out, witness:

- (7) Paul maakt de laatste tijd een gespannen indruk.
 “Recently, Paul seems to be under stress”
- a. Misschien komt dat omdat hij zelden de kat aait
 maybe comes that because he seldom the cat pets
 - b. Misschien komt dat omdat hij de kat zelden aait
 maybe comes that because he the cat seldom pets
 “That’s maybe because he hardly ever pets the cat”

In (7) no cat is introduced in the preceding discourse, hence no linguistic antecedent is provided for an anaphoric interpretation of the cat. Yet, the scrambled variant (7b) is not excluded in this context. And moreover, scrambling does not force an anaphoric interpretation in this context. In fact, the actual context makes the non-anaphoric interpretation for the cat the preferred one in both (7a) and (7b). Similarly, the context in (8) where an antecedent is introduced for the definite *the cat*, triggers the anaphoric reading in both (8a) and (8b):

- (8) Paul heeft een kat die de laatste tijd een gespannen indruk maakt.
 “Paul has a cat that seems to be under stress, recently”
- a. Misschien komt dat omdat Paul zelden de kat aait
 maybe comes that because Paul seldom the cat pets
 - b. Misschien komt dat omdat Paul de kat zelden aait
 maybe comes that because Paul the cat seldom pets
 “That’s maybe because Paul hardly ever pets the cat”

I conclude that when there are two interpretations possible (e.g., *the cat* can be interpreted anaphorically or non-anaphorically), then in the absence of further contextual clues, word order indicates which interpretation is the preferred one ((6)). Yet, the presence of an actual context (in (7) and (8)) can easily overrule this word order effect. This leads me to the following generalization:

- *Generalization*

Adding a specific context decreases the number of possible interpretations and (therefore) increases the number of word order possibilities.

In other words, if one interpretation is pragmatically preferred (in the above cases by adding a specific context), then the effect of word order on interpretation becomes negligible. I will provide two more arguments that support this view.

First, there are definites that preferably get a non-anaphoric interpretation anyway. An example of such a definite is the one in (4) above: *de was* ‘the laundry’ is usually interpreted non-anaphorically. Similarly, *de koningin* ‘the queen’ in (9) usually refers independently of the discourse to the one and only queen of the Netherlands in a language like Dutch:

- (9) a. omdat ik gisteren de koningin zag
 since I yesterday the queen saw

- b. omdat ik de koningin gisteren zag
since I the queen yesterday saw
“since I saw the queen yesterday”

Scrambling the definite in (9) is completely optional; there is no difference in interpretation nor in markedness and there is no tendency to interpret the scrambled definite as anaphoric in the absence of any context. On the basis of examples such as (4) and (9) we conclude that if one interpretation is pragmatically preferred (in these cases the non-anaphoric interpretation for *the laundry* and *the queen*), then word order becomes negligible.

The same can be demonstrated with respect to scope ambiguities:

- (10) a. dat Jan drie keer alle jongens kuste
that Jan three times all boys kissed
“Three times, Jan kissed all boys”
- b. dat Jan alle jongens drie keer kuste
that Jan all boys three times kissed
“All boys, Jan kissed three times”

The pattern in (10) indicates that the preferred interpretation for the relation between two quantifiers is the interpretation that is directly reflected by the word order (cf. Diesing and Jelinek 1995, Ruys 1996). The quantifying noun phrase *all boys* is in the scope of the adverbial quantifier *drie keer* in (10a) and vice versa in (10b).

But when there is in fact no scope ambiguity because only one interpretation is pragmatically obvious, the word order that is not in accordance with the intended interpretation is however possible and it does not evoke the other (non-preferred) interpretation:

- (11) a. dat Jan minstens een keer per jaar al z'n vrienden bezoekt (maar bij
that Jan at least once a year all his friends visits (but by
voorkeur niet op hun verjaardag)
preference not on their birthday)
“All his friends, Jan visits at least once a year (but preferably not on their birthday)”
- b. dat Jan alle brieven een keer verscheurd heeft
that Jan all letters once torn up has
“Once, Jan torn up all letters”

Again, the conclusion must be that if one interpretation is pragmatically preferred (in these cases the scope order as given by the translation), then word order becomes negligible.

Thus, in general, the following conclusion can be formulated:

- *Conclusion*

If there is a conflict in Dutch between the interpretation favoured by context and the interpretation favoured by word order, context wins.

This can be illustrated once more with respect to the following two examples, adapted from Neeleman and Reinhart (1998). Neeleman and Reinhart claim that (12) is an example of a context that favours scrambling, whereas (13) is an example of a context

that disfavors scrambling. In fact, however, (12) and (13) are perfectly well-formed and show the regular intonation patterns that fit in the context, which are equivalent to the intonation patterns of their (un)scrambled counterparts (basically, the new information is focused, but not in an extra-ordinary way).

- (12) A: Hoe gaat het met de review van Jans boek?
 how goes it with the review of Jan's book
 B: Nou, ik heb eindelijk het boek gelezen, maar ik snap er nog niet
 now I have finally the book read but I understand there yet not
 veel van
 much of
- (13) A: Heeft je buurman gisteren de deur geverfd?
 has your neighbour yesterday the door painted
 B: Nee, maar hij heeft de dakgoot gisteren wel geverfd
 no but he has the gutter yesterday AFF painted
 "No, but he did paint the gutter yesterday"

Neeleman and Reinhart provide the answer of speaker B in (12) without the clarifying *but*-clause, which makes it rather odd as an answer to the question, independently of scrambling. In (13), they use the definite *the window*, which makes the sentence slightly odd (albeit not ill-formed, in my opinion) because of the uniqueness condition on definites, that is automatically satisfied by definites such as *the door*, *the gutter* or the plural *the windows*, but not as easily by *the window*. Again, this does not depend on scrambling. In fact, the definite *de deur* 'the door' can already scramble in the question without causing an odd or ill-formed question: *Heeft je buurman de deur gisteren geverfd?* 'has your neighbour the door yesterday painted'.

Neeleman and Reinhart claim that in unscrambled VPs, the default sentence stress falls on the object, whereas in scrambled versions it falls on the verb. The focus set of a sentence consists of all and only the constituents which contain the default stress. Therefore, the focus set of an unscrambled sentence is {IP,VP,O} (that is, it includes the object, but not the verb alone), whereas the focus set of a scrambled structure is {IP,VP,V} (which includes the verb, but not the object). In Dutch, then, a scrambled object is not in a position to be assigned default stress. Hence, it can be used only if it is appropriate for the object to be fully destressed. Neeleman and Reinhart, following Neeleman's (1994) base-generation account of scrambling, argue that scrambling is not costly at all (there can be no economy difference related to where we choose to place the adverb), and hence, less costly than the use of marked stress. Thus, they predict that if a context requires the verb to be the focus of a sentence, then in a language like Dutch, scrambling is obligatory. Likewise, if a context requires the object to be in focus, scrambling is not allowed. Therefore, (12) and (13) should be inappropriate or ill-formed. In (13) the context set solely selects the object as the focus and that is why the scrambled variant should not be used in this context. Yet, (12) and (13) are not as bad as they should be; in fact, they are perfect.

2. Optimal scrambling

In Optimality Theory (Prince and Smolensky, to appear, 1997) a grammar consists of a set of well-formedness constraints. These constraints apply simultaneously to representations of structures and they are soft, which means potentially violable and typically conflicting. Furthermore, at least an important subset of these constraints is shared by all languages, forming part of Universal Grammar. Individual languages rank these universal constraints differently in their language-specific hierarchies in such a way that higher ranked constraints have total dominance over lower ranked constraints. Possible output candidates for each underlying form are evaluated by means of these constraint rankings. The output that best satisfies the constraints is the optimal candidate and will be the realized form. By analyzing the results arising from ranking the universal constraints in all possible dominance hierarchies, one can predict and explain which surface patterns are possible in natural languages.

Choi (1996) provides an Optimality Theory (henceforth, OT) approach to scrambling. She argues that there are in fact two discourse notions involved in the process of scrambling. Elements which are [-new] scramble, but also elements which are [+prominent] scramble, two constraints which do not only capture scrambling of topical/anaphoric elements, but also of elements which are contrastively focused:

- (14) Ik heb het BOEK gisteren gelezen, niet de KRANT
I have the book yesterday read, not the newspaper
“I read the BOOK yesterday, not the NEWSPAPER”

The information structuring constraints Choi uses to account for scrambling phenomena are the following:

- (15) a. *NEW: A [-new] element should precede a [+new] element*
b. *PROM: A [+prominent] element should precede a [-prominent] element*

In Choi's analysis, using discourse features ([+/- new] and [+/- prominent]) as part of the input representation, each scrambled variant is the best structural description of a particular information structure with respect to a small number of syntactic and discourse constraints. In a language like Dutch or German the prominence constraint (15b) outranks (15a): PROM >> NEW. Choi correctly observes that definites do not obligatorily scramble, not even when they bear the feature [-new]. In Choi's OT approach, optional scrambling of definites is explained by the interaction of the informational constraint NEW and CN2, a syntactic constraint that favours canonical (unscrambled) word order.²

Crucially in Choi's analysis, NEW and CN2 are not ranked with respect to each other. Because NEW and CN2 are equally strong, syntactic structures with (un)scrambled

² We could also opt for a constraint such as Grimshaw's (1997) *STAY* here (cf. Costa 1998, Merchant 1997). I prefer a constraint such as Choi's *CN2* because it is compatible with the view that scrambling involves base-generated structures.

definites can both be well-formed. This approach cannot account for all cases of optional scrambling, however. For instance, topics are characterized as [+prominent] and [-new], hence they would violate both information structuring constraints in unscrambled position. In other words, topics should scramble obligatorily. We have seen that this predicted obligatoriness is not in accordance with the facts, however. For example, Paul's cat can be considered a [-new],[+prominent] topic in (8), but it does not obligatorily scramble. At the same time, Choi cannot account for the fact that a [+new], [-prominent] definite such as *the laundry* in (4) does scramble, thereby violating CN2, while NEW and PROM don't apply. In other words, Choi's theory cannot account for the fact that scrambling is truly optional for definites irrespective of their discourse features.

There is one other obvious problem for an approach such as Choi's, and that is the different behaviour of the definites and indefinites with respect to scrambling in (2) and (4), irrespective of their similar [+new], [-prominent] discourse status. For Choi, syntactic and semantic information (i.e., the (in)definiteness of the noun phrases under consideration) is totally ignored in favour of discourse information. The fact, however, that [+new],[−prominent] definites such as the one in (4) above optionally scramble, just like [−new][+prominent] definites such as the one in (3), but unlike [+new][−prominent] indefinites as in (2), indicates that a proper analysis of scrambling should not entirely be based on discourse features reflecting topic-focus structure or prominence.³

In Van der Does and De Hoop (1998) it is argued that the difference between indefinite and definite noun phrases with respect to optional scrambling, follows from a difference in their semantic type. The weak or predicative reading of indefinites (such as in (2) above) is argued to follow from the semantic, predicative type these indefinites naturally have. These predicates are incorporated by an incorporating verb type. Definites like those in (4) can intuitively be characterized as predicative, like their indefinite counterparts, but they are independent of the verb for their interpretation. That is, definites can be independent of their semantic context in a way indefinites cannot and they denote most naturally unique objects of the kind the nominal indicates. This explains why definites scramble freely: merging an incorporating verb and a predicative definite turns out to be equivalent to merging an ordinary transitive type of verb and an ordinary referential type of definite. Definites, even weak ones such as in (4), do not have to denote in type $\langle e, t \rangle$ in order to allow for an interpretation that is equal to the interpretation yielded by a process of semantic incorporation. Following the strategy in Partee and Rooth (1983) to interpret an NP in as simple a type as possible, definite NPs have their basic denotation in type e (the referential type) whereas indefinite NPs denote most naturally in type $\langle e, t \rangle$ (the predicative type).

³ A connection between specificity and information status is indirectly responsible for the limited scrambling options of indefinite NPs, according to Choi. A nonspecific NP is informationally dependent on the bigger information unit. In Choi's theory scrambling is only possible when the phrase has an independent information status. Intuitively, however, a definite such as *the laundry* in (4), is informationally equivalent to the informationally dependent *a piss* in (2). Unfortunately, Choi does not discuss this type of definite.

- (16) *Basic interpretation: interpret an NP in as simple a type as possible, hence:*
- a. *Interpret a definite in type e*
 - b. *Interpret an indefinite in type (e,t)*
 - c. *Interpret a quantifier in type $((e,t),t)$*

While predicative indefinites are dependent on the predicate for their existential interpretation, definites are not. The *iota*-operation that can be used to shift the type of a predicative definite in type (e,t) to type e is not dependent on the verb. This explains why definites scramble more freely than indefinites. For the technical details of the analysis, I refer to Van der Does and De Hoop (1998).

In other words, from a computational point of view, scrambling of definites is truly optional. Yet, one would also like to account for the interpretive tendencies that arise with scrambling which led previous authors to assume a mapping between structure and discourse. In this paper I will account for these tendencies by means of an unordered constraint set along the lines of Anttila and Yu Cho (1998).

3. Producing scrambling

In fact, it is well-known that definites scramble more often than indefinites. This type of frequency information appears to be of utmost importance when we judge the relative well-formedness or markedness of (un)scrambled sentences, perhaps more relevant than the actual discourse properties of the (un)scrambled noun phrases in the context. That is, the fact that the scrambled variant *dat ik de was nog moet doen* 'that I still have to do the laundry' is perfectly well-formed and unmarked, although *de was* is neither a topic nor contrastively focused, indicates that its definite shape is already sufficient to make the scrambled structure unmarked. Indefinites that scramble, will in any case lead to structures with a higher degree of markedness.⁴

⁴ Neeleman and Reinhart's (1998) approach has the same shortcoming as Choi's in this respect; they do not recognize the intrinsic difference between definites on the one hand and indefinites on the other. Whereas Choi tries to reduce definiteness effects of scrambling to discourse properties such as newness and prominence, Neeleman and Reinhart try to reduce these effects to default intonation patterns. Anaphoric expressions are typically destressed and that is how Neeleman and Reinhart derive the definiteness effect of scrambling (the fact that definites scramble more often than indefinites is then related to the fact that they are more typically anaphoric). According to Neeleman and Reinhart, scrambling is appropriate only in a context which enables full destressing of the object, and the most typical context allowing that is that of anaphoric NPs, and most typically, definite, but not indefinite NPs can be anaphoric. It is well-known that definites scramble not as often as pronouns and more often than indefinites and this may be related to their discourse and/or phonological properties. That does not alter the fact, however, that this type of frequency information itself influences our judgement of the relative well-formedness or markedness of (un)scrambled sentences, partly independent of the actual contexts of utterance. This explains that while scrambled indefinites are only considered well-formed when certain conditions are met, scrambled definites lead to well-formed sentences irrespective of their discourse status. That is, the

In OT different families of constraints are distinguished. One class of universal constraints in OT formalizes the notion of markedness. These constraints prefer unmarked structures. A second class of universal constraints in OT is constituted by the so-called faithfulness or correspondence constraints. For example, faithfulness or correspondence constraints tie the success of an output candidate to the shape of the corresponding input. Faithfulness or correspondence constraints ensure that not too many lexical distinctions are wiped out by the markedness constraints. There are also correspondence constraints that indicate the mutual relation between output forms (output-output correspondence). These constraints compare candidate syntactic surface forms not to an underlying structure but to the surface form of paradigmatically related words. Correspondence constraints reflect for example ‘cyclic’ effects in phonology (cf. Burzio 1995, 1998).

What is missing in Choi’s (1996) analysis, and what we need to account for the unmarkedness of scrambled definites irrespective of their discourse features is a family of constraints that deals with output-output correspondence in the case of scrambling. That is, it appears that the well-formedness or markedness of (un)scrambled structures partly depends on their degree of faithfulness to the output structures they correspond to. The type of constraint that interacts with NEW and PROM is not (only) a structural constraint that favours canonical (i.e., unscrambled) word order. Indeed, what counts as canonical word order heavily depends on the type of noun phrase involved. Therefore, I propose to use three correspondence constraints as in (17).

- (17) *Surface correspondence:*
- a. SC1: *Definites scramble*
 - b. SC2: *Indefinites don’t scramble*
 - c. SC3: *Pronouns scramble*

Two other constraints are borrowed from Choi in a slightly different formulation:

- (18) *NEW (new version): Anaphors scramble*⁵

fact that scrambling in (4) is perfectly well-formed and unmarked, although *de was* is neither anaphorically destressed nor contrastively focused, indicates that its definite shape is already sufficient to make the scrambled structure unmarked, irrespective of the actual intonation pattern.

⁵ This relation might be an indirect relation in Dutch, mediated by default sentence accentuation, as proposed by Neeleman and Reinhart, a.o. Default or unmarked sentence accent is strongly related to syntactic structure (cf. Cinque 1993, Neeleman and Reinhart 1998, Merchant 1997), which can be captured by (ia). This constraint can replace Choi’s NEW and together with (ib) it will ensure that anaphors scramble. In (ib,c) two constraints capture the relation between topic and focus and accentuation (cf. Terken and Nootboom 1987).

- (i)
- a. *NEW: Main sentential pitch accent falls on the most deeply embedded (lexical) constituent*
 - b. *Topics (anaphors) are deaccented*
 - c. *Focused elements are accented*

Preferably, NEW as formulated in (18) will go hand in hand with deaccenting, although there are indications that when there are conflicts between context and accent, context wins (a.o., Van Donselaar 1995).

(19) *CN2: Favour unscrambled word order*

I propose that SC1, NEW, and CN2 are not ranked with respect to each other. That means that we obtain a set of six total orders. The number of rankings is inversely proportionate to the number of tableaux: the fewer rankings, the more tableaux. Anttila and Yu Cho (1998) account for optionality and statistical preferences in the following way:

- (20) a. An output candidate is predicted by the grammar iff it wins in some tableau.
 b. If a candidate wins in n tableaux and t is the total number of tableaux, then the candidate's probability of occurrence is n/t .

Let's consider an input with an anaphoric definite first (examples like (3) and (8) above):

(21) *Constraint tableau*

Input (3), (8)	Output	SC1	NEW	CN2
anaphoric	+ scrambling			*
definite	- scrambling	*	*	

As pointed out above, I propose that the three constraints involved are not ranked with respect to each other. That means that both output candidates are optimal under certain rankings of the constraints. Hence, both candidates will occur in Dutch, yet not equally frequently. This is illustrated in the tableau in (22), which contains all six possible total rankings in the unordered set of constraints {SC1, NEW, CN2}:

(22) *Results of possible total rankings*

Total orders {SC1, NEW, CN2}	Winning candidate in (21)	Result
SC1 >> NEW >> CN2	+ scrambling	anaphoric definites scramble in 2/3 of the cases
SC1 >> CN2 >> NEW	+ scrambling	
NEW >> SC1 >> CN2	+ scrambling	
NEW >> CN2 >> SC1	+ scrambling	
CN2 >> SC1 >> NEW	- scrambling	anaphoric definites don't scramble in 1/3 of the cases
CN2 >> NEW >> SC1	- scrambling	

The analysis proposed here accounts for the fact that even if it is not true that anaphors obligatorily scramble, we do scramble them more often than we leave them in situ. Let us now consider non-anaphoric definites, such as in (4), (7), and (9) above:

(23) *Constraint tableau*

Input (4), (7), (9)	Output	SC1	NEW	CN2
non-anaphoric definite	+ scrambling			*
	- scrambling	*		

(24) *Results of possible total rankings*

Total orders {SC1, NEW, CN2}	Winning candidate in (23)	Result
SC1 >> NEW >> CN2	+ scrambling	non-anaphoric definites scramble in 1/2 of the cases
SC1 >> CN2 >> NEW	+ scrambling	
NEW >> SC1 >> CN2	+ scrambling	
NEW >> CN2 >> SC1	- scrambling	non-anaphoric definites don't scramble in 1/2 of the cases
CN2 >> SC1 >> NEW	- scrambling	
CN2 >> NEW >> SC1	- scrambling	

NEW is vacuously satisfied in the case of non-anaphoric definites. As a consequence, the tableau in (24) nicely conforms to the intuition that scrambling of a non-anaphoric definite such as *de koningin* 'the queen' in (9) above is truly optional. That is, the conflict between the output-output correspondence constraint SC1 ("Definites scramble") and the markedness constraint CN2 ("Don't scramble") predicts scrambling of such a definite in 50% of the cases.

4. Interpreting scrambling

Let us now return to the interpretation of (un)scrambled definites. We accounted for the fact that despite the constraint NEW in (18), the cat in (8a) does not have to scramble in order to get interpreted as an anaphor linked to the antecedent *a cat* in the preceding sentence. Note also that it is in fact impossible to interpret the cat in (8) as non-anaphoric (which means that it cannot refer to a cat other than the one introduced by the linguistic antecedent; for instance, it cannot refer to the speaker's cat). There is a pragmatic constraint, which may be formulated in very general terms (such as 'Be relevant' or 'Be coherent'), but for which I will use a more specific instantiation in this context, namely Williams' (1997) *DOAP*:

(25) *DOAP: Don't Overlook Anaphoric Possibilities:
Opportunities to anaphorize text must be seized*

This contextual constraint turns out to be stronger than the constraint NEW in (18). Interpreting a sentence is not the same as producing or even judging a sentence (cf. Prince and Smolensky 1997, and Smolensky 1996, on the relation between production

and interpretation of language). Dutch children produce scrambled as well as unscrambled definites as soon as they also start using adverbial elements (Neeleman and Weerman 1997). Neeleman and Weerman furthermore cite Schaeffer (1997) who shows that from the very first stages the syntactic option of scrambling is present in Dutch children. Not only do two-year-olds have the option of scrambling, there is also a development in the usage of scrambling. Schaeffer concludes that this development is due to a discourse principle concerning specificity, which has not yet been acquired by the two-year-olds, but which is present in the three-year-olds. However, if this were a proper explanation, one would expect that two-year-olds scramble indefinites as often as definites, which is — I bet — contrary to fact.

In this light, note the striking results from two psycholinguistic experiments concerning children's interpretation of (un)scrambled indefinites in Dutch (Krämer 1998). It turns out that almost all children until the age of 8 interpret scrambled indefinites as well as unscrambled ones like predicates (the most natural semantic type for indefinites). That is, they do not take into consideration word order clues for the interpretation of indefinites but rather ignore word order and interpret the indefinite objects in the most unmarked way, i.e., in type *(e,t)* (which is in accordance with the constraint family formulated in (16) above).

We have seen that scrambling of definites is optional. The idea that scrambling is optional is perfectly compatible with a syntactic base-generation approach of scrambled and unscrambled structures (Neeleman 1994). We cannot ignore, however, the interpretive tendencies that arise with scrambling. Whereas Neeleman and Reinhart (1998) and Choi (1996) both argue that the context in which a sentence is used imposes restrictions on which word orders may be chosen, I made the observation that context can easily overrule the interpretive effects that follow from word order. If contextual or lexical clues do not bias a certain interpretation, then word order may be decisive in what a hearer/reader takes to be the intended reading. If context and word order both point into the same direction, there is no conflict and the structure is unmarked in this respect. But if there is a conflict between the interpretation evoked by the context and the one that should be in accordance with the given word order, then context wins.

This is exactly what strict domination hierarchies in OT provide: no amount of success on weaker constraints can compensate for failure on a stronger one. This corresponds to the numerical strength of a constraint being so much greater than the strengths of those constraints ranked lower in the hierarchy that the combined force of the lower-ranked constraints cannot exceed the force of the higher-ranked constraint. I claim that in the interpretation of (well-formed) structures, the contextual constraint DOAP outranks the constraint NEW. That is why the interpretation of so-called marked word orders is not a problem at all as long as there is no ambiguity that has to be resolved and for which the chosen word order could be used as a clue.

Note that while OT syntax maps a semantic input onto a(n) (set of) optimal syntactic output(s), OT semantics maps a syntactically well-formed input onto a (set of) optimal interpretation(s) (cf. Hendriks and De Hoop 1998). In the previous section the OT syntax tableaux reflected the mapping of (non-)anaphoric definites to (un)scrambled structure. At this point we illustrate the mapping of (un)scrambled definites to (non-)anaphoric

interpretations by means of an OT semantics tableau. The reader may verify that pure word order constraints such as SC1 and CN2 do not play a part here (both are either violated or satisfied by all relevant output candidates). Likewise, in the OT syntax tableaux of the previous section, the interpretive constraint DOAP is vacuously satisfied.

(26) *OT semantics constraint tableau*

Input	Output	DOAP	NEW
linguistic antecedent + scrambled definite	anaphoric interpretation		
	non-anaphoric interpretation	*	
linguistic antecedent + unscrambled definite	anaphoric interpretation		*
	non-anaphoric interpretation	*	

Conclusion

The interdependence between word order, default accentuation, and information structure is only used in the interpretation of (un)scrambled syntactic structures in the absence of further context and actual accentuation. Contextual information as well as actual accentuation can result in the activation of a certain interpretation despite the violation of constraints on the relation between word order and interpretation. Furthermore, frequency of scrambling of certain types of noun phrases has a considerable impact on the production of scrambling, partly independent of discourse interpretation. In this paper I provided an Optimality Theoretic account of the interpretive tendencies that are associated with scrambling of definites in Dutch.

References

Anttila, A. and Y.-m. Yu Cho (1998) ‘Variation and Change in Optimality Theory’, *Lingua* 104, 31–56.
 Burzio, L. (1995) ‘The Rise of Optimality Theory’, *Glott International* 1–6.
 Burzio, L. (1998) ‘Multiple Correspondence’. *Lingua* 104, 79–110.
 Butt, M. and T. King (1996) ‘Focus, Adjacency, and Nonspecificity’. Unpublished manuscript. Stuttgart/Stanford.
 Choi, H.-W. (1996) *Optimizing Structure in Context: Scrambling and Information Structure*. PhD dissertation. Stanford University.
 Cinque, G. (1993) ‘A null theory of phrase and compound stress’, *Linguistic Inquiry* 24, 239–297.
 Costa, J. (1998) *Word Order Variation. A constraint-based approach*. PhD dissertation. Leiden University.

- Does, J. van der and H. de Hoop (1998) 'Type-Shifting and Scrambled Definites'. *Journal of Semantics* 15, 393–416.
- Diesing, M. and E. Jelinek (1995) 'Distributing Arguments', *Natural Language Semantics* 3, 123–176.
- Donselaar, W. A. van (1995) *Effects of Accentuation and Given/New Information on Word Processing*. Ph.D. dissertation. Utrecht University.
- Geenhoven, V. Van (1996) *Semantic Incorporation and Indefinite Descriptions. Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. PhD dissertation. Tübingen.
- Grimshaw, J. (1997) 'Projection, Heads, and Optimality', *Linguistic Inquiry* 28, 373–422.
- Hendriks, P. and H. de Hoop (1998) 'On the interpretation of semantic relations in the absence of syntactic structure'. Unpublished manuscript. Groningen/Utrecht.
- Hoop, H. de (1992) *Case Configuration and Noun Phrase Interpretation*. PhD. dissertation. University of Groningen. (1996, Garland Publishing, New York & London).
- Krämer, I. (1998) 'Children's Interpretations of Indefinite Object Noun Phrases'. In R. van Bezooijen and R. Kager (eds), *Linguistics in the Netherlands 1998*. John Benjamins, Amsterdam/Philadelphia.
- Merchant, J. (1997) 'Anaphoric destressing and scrambling in Dutch and English'. Paper presented at the Hopkins OT Workshop. Johns Hopkins University.
- Neeleman, A. (1994) 'Scrambling as a D-structure Phenomenon'. In N. Corver and H. van Riemsdijk (eds), *Studies on Scrambling; Movement and Non-Movement Approaches to Free Word Order Phenomena*. Mouton de Gruyter, Berlin.
- Neeleman, A. and T. Reinhart (1998) 'Scrambling and the PF Interface'. In M. Butt and W. Gueder (eds), *The Projection of Arguments. Lexical and Compositional Factors*. CSLI, Stanford.
- Neeleman, A. and F. Weerman. 'L1 and L2 Word Order Acquisition', *Language Acquisition* 6, 125–170.
- Partee, B. H. and M. Rooth (1983) 'Generalized Conjunction and Type Ambiguity'. In R. Bäuerle, C. Schwarze and A. von Stechow (eds), *Meaning, Use and Interpretation of Language*. De Gruyter, Berlin.
- Prince, A. and P. Smolensky (to appear) *Optimality Theory: constraint interaction in generative grammar*. Unpublished Manuscript, 1993. To appear in Linguistic Inquiry Monograph Series. MIT Press, Cambridge.
- Prince, A. and P. Smolensky (1997) 'Optimality Theory: From Neural Networks to Universal Grammar', *Science* 275, 1604–1610.
- Ruys, E. (1996) 'Dutch Scrambling and the Strong-Weak Distinction'. Unpublished manuscript. OTS, Utrecht.
- Schaeffer, J. (1997) *Direct Object Scrambling in Dutch and Italian Child Language*. PhD dissertation. UC Los Angeles.
- Smolensky, P. (1996) 'On the comprehension/production dilemma in child language', *Linguistic Inquiry* 27, 720–732.
- Terken, J. M. B. and S. G. Nootboom (1987) 'Opposite effects of accentuation and deaccentuation on verification latencies for "given" and "new" information', *Language and Cognitive processes* 2, 145–163.
- Williams, E. (1997) 'Blocking and Anaphora', *Linguistic Inquiry* 28, 577–628.
- Zwart, J.-W. (1995) 'Word Order, Intonation, and Noun Phrase Interpretation in Dutch', *WECOL Proceedings*.