

Scrambling of definite object NPs in Dutch¹

Formal theories, corpus data and experimental research

Hans Broekhuis (Meertens Institute, Amsterdam)
March 16, 2021

1 Introduction

Nominal arguments often occupy different positions in the so-called middle field of the Dutch clause, that is, the part of the clause bounded to the left by the C-position filled by the complementizer in embedded clauses and by the finite verb in main clauses and bounded to the right by the verbs (if present) in clause-final position. The literature on this variation in word order focuses specifically on the position of nominal direct objects in relation to specific kinds of clause adverb, especially those expressing propositional modality such as *waarschijnlijk* ‘probably’. We illustrate this in (1a) by means of an embedded clause; the use of angled brackets indicate alternative placements of the nominal object. Note that a similar word-order variation is also found with subjects, as is shown in (1b).

- (1) • Scrambling
- a. dat de buurman <zijn huis> waarschijnlijk <zijn huis> verkoopt.
 that the neighbor his house probably sells
 ‘that the neighbor will probably sell his house.’
- b. dat <de buurman> waarschijnlijk <de buurman> zijn huis verkoopt.
 that the neighbor probably his house sells
 ‘that the neighbor will probably sell his house.’

The current standard assumption in generative grammar is that the object and the subject are both base-generated within the lexical projection of the main verb: the object is taken to be the complement of a lexical root element *V*, while the subject is the specifier of a light verb *v* carrying the verbal specifications of the root element. The base structure is therefore as follows: [_{VP} SU *v* [_{VP} OBJ *V*]].² In order to account for the scrambled word orders, it is assumed that the nominal arguments are “optionally” moved into positions to the left of the default position of the relevant kinds of clause adverb, where they can be assigned (abstract) case. If scrambling were truly optional in Dutch, we would predict all output structures in (2a-e) to be acceptable; ADV_s stands for the set of adverbs demarcating the boundary of the lexical projection of the verb (*vP*). It is important to note that the movements in (2) may also occur when ADV_s is absent, although they will then often have no effect on the linear order of the clause.

¹ hans.broekhuis@meertens.knaw.nl. I like to thank Gert-Jan Schoenmakers for discussing the material in this article with me and Frits Beukema for correcting my English: their comments have led to various improvements; all remaining errors are mine.

² When I discuss works adopting the earlier structure [_{VP} SU [_{V'} OBJ *V*]], the terminology will be adapted to the more recent insight: *vP* and *VP* will then be used instead of, respectively, *VP* and *V'*. For the sake of simplicity, the structures in (2) are given in the default order of the main verb and its nominal arguments in Dutch embedded clauses (SOV). There are good reasons, however, for assuming that word order is not determined within syntax but involves linearization of the output structures of syntax in the phonological component of the grammar: see Vanden Wyngaerd (1988/1989), Kayne (1994), Chomsky (1995), and Zwart (1993/1997) for relevant discussion.

- (2) • Scrambling
- | | | |
|----|---|-----------------------------------|
| a. | [... C ... ADV _S [_{VP} SU v [_{VP} OBJ V]]]. | [(1b) without subject scrambling] |
| b. | [... C SU _j ... ADV _S [_{VP} t _j v [_{VP} OBJ _i V]]]. | [(1b) with subject scrambling] |
| c. | [... C SU _j ... ADV _S [_{VP} t _j v [_{VP} OBJ _i V]]]. | [(1a) without object scrambling] |
| d. | [... C SU _j OBJ _i ADV _S [_{VP} t _j v [_{VP} t _i V]]]. | [(1a) with object scrambling] |
| e. | *[... C ... OBJ _i ADV _S [_{VP} SU v [_{VP} t _i V]]]. | [unacceptable in Dutch]. |

At first sight, the assumption that scrambling is optional in Dutch seems to be close to the mark. It cannot be fully correct, however, given that the order in (2e) is impossible in Dutch: cf. **dat het huis_i waarschijnlijk de buurman t_i verkoopt*. The unacceptability of this example can be accounted for by assuming that scrambling is subject to a restriction holding that the neutral order of the nominal arguments (S<IO<DO) cannot be changed. This is not the only restriction on scrambling: the examples in (3) show that the availability of object scrambling also depends on the type of NP. Non-specific indefinite (pro)nominal objects such as *een boek* in (3a) cannot be scrambled, definite nominal objects such as *het boek* are “optionally” scrambled, while definite pronominal objects like *het* ‘it’ must be scrambled. Note that, for the sake of brevity, we will use the abbreviation NP for lexical (non-pronominal) noun phrases like *een/het boek* from now on, although the notion of noun phrase also includes pronouns.

- (3)
- | | | |
|----|--|------------------------------|
| a. | dat Jan <*een boek> waarschijnlijk <een boek> koopt. | [non-specific indefinite NP] |
| | that Jan a book probably buys | |
| b. | dat Jan <het boek> waarschijnlijk <het boek> koopt. | [definite NP] |
| | that Jan the book probably buys | |
| c. | dat Jan <het> waarschijnlijk <*het> koopt. | [definite pronoun] |
| | that Jan it probably buys | |

The examples in (4) show that the same holds for subjects; we will not discuss the complication that non-specific indefinite (pro)nominal subjects normally require the expletive *er* ‘there’ to be present in monadic constructions. We will in fact more or less ignore subject scrambling from now on, as it does not play an important role in the remainder of this article, and refer the reader to Broekhuis (2007) for detailed discussion.

- (4)
- | | | |
|----|---|------------------------------|
| a. | dat er <*een baby> waarschijnlijk <een baby> huult. | [non-specific indefinite NP] |
| | that there a baby probably cries | |
| b. | dat <de baby> waarschijnlijk <de baby> huult. | [definite NP] |
| | that the baby probably cries | |
| c. | dat <hij/zij> waarschijnlijk <*hij/*zij> huult. | [definite pronoun] |
| | that he/she probably cries | |

The pattern in (3) has given rise to a long-standing consensus that object scrambling is typically related to the information structure of the clause: canonically scrambled noun phrases belong to the PRESUPPOSITION (also called theme or “old” information) of the clause, whereas noun phrases that are not scrambled are typically part of the FOCUS (also rheme, “new” information or comment) of the clause. See, especially, Verhagen (1986) and references cited there, as well as Neeleman & Reinhardt (1998), who formulate the relevant generalization (less accurately in my view) in terms of anaphoricity.³ This claim provides a simple account for the pattern in (3):

³ Presuppositionality involves activated information from the common ground (the information shared by speaker and addressee) regardless of the origin of that information, while anaphoricity involves “earlier mention” of entities in the discourse. The former notion is both narrower and wider than the latter: narrower because earlier mentioned information may become deactivated in the course of the discourse, and wider because information not explicitly mentioned can be presuppositional because it is inferable from the discourse, established in earlier discourse, well-known, etc. The literature reviewed below generally refers to anaphoricity.

nonspecific indefinite NPs typically belong to the focus of the clause, for which reason they cannot be scrambled, while definite pronouns typically belong to the presupposition of the clause, for which reason they must be scrambled; definite NPs, on the other hand, can be part of either the focus or the presupposition of the clause and consequently undergo “seemingly” optional scrambling. Most of this have appeared to be uncontroversial up to till now but for one thing: Van der Does & De Hoop (1998) suggested that object scrambling of definite NPs is “truly” optional in the sense that it is independent of the information structure of the clause. This claim has subsequently been investigated in various corpus/experimental studies: although the results of these studies do not seem to confirm the original claim, they have led to the surprising conclusion that scrambling of definite object NPs is very infrequent in speech; cf. Van Bergen & De Swart (2009/2010/2011), Schoenmakers & De Swart (2019) and Schoenmakers (2020). De Hoop (2016) has further used the studies by Van Bergen & De Swart for arguing against the validity of the earlier studies based on introspection data. Although Broekhuis (2016) already briefly countered De Hoop’s arguments, her claim justifies a more extensive evaluation of the corpus/experimental studies in order to see whether they indeed show that the findings of earlier theoretical research are flawed. In order to be able to do this, I will first provide more background on the notion of scrambling, and subsequently discuss the body of work mentioned above in chronological order.

2 Background information on A-scrambling

The notion of scrambling is a rather unspecific one and seems to cast the net too wide. It was originally introduced to account for “free word order” languages: Latin, for instance, allows clausemates (the clausal constituents—including the verb—making up a single clause) to surface in various orders, and (in poetry at least) clausal constituents to be split. Ross (1967: §3.2.1) characterizes scrambling as a transformational rule in the stylistic component of the grammar that optionally inverts the major constituents of a clause. From there on, the term has been used in the description of various language types, often to refer to completely different word-order phenomena. It is therefore not surprising that the theoretical accounts of “scrambling” are disparate: according to some, scrambling involves A-movement (movement of a nominal argument licensed by agreement and case features) while others maintain that it is A'-movement (movement that is not restricted to nominal arguments and licensed by other—often semantic—features). This confusing picture also surfaces in the discussion of Dutch/German scrambling, which has been analyzed as A-movement, A'-movement, and even as movement with mixed A and A'-properties; see the introductions and the individual studies in Corver & Van Riemsdijk (1994) and Karimi (2003) for detailed reviews and ample illustration.

Neeleman (1994a/1994b) provides a first, crucial step in clarifying this confusing picture by showing that Dutch has in fact two *distinct* kinds of scrambling: A-scrambling, which was the main topic of earlier studies and which was already briefly introduced in (1) and (3) above, and A'-scrambling.⁴ The latter variety is not restricted to nominal arguments but can also be applied to non-arguments: there are, for instance, rules of contrastive topic/focus and negation movement, which move contrastive topics/foci and negative phrases into a more leftward position in the middle field of the clause. Negation movement is illustrated in (5) by means of the PP-complement of the adjective *trots* ‘proud’: while the PP is normally located behind the adjective, it obligatorily moves into a more leftward position when it expresses sentence negation; cf. Haegeman (1995: 179).

⁴ Neeleman proposed a base-generation account of A-scrambling instead of the A-movement analysis given in (3). I refer to Broekhuis & Corver (2016: 1611-2) for a discussion of some empirical problems for the base-generation approach.

- (5) a. dat Marie waarschijnlijk [_{AP} erg dol [_{PP} op Peter]] is. [neutral order]
 that Marie probably very fond of Peter is
 ‘that Marie is probably very fond of Peter.’
 b. dat Marie waarschijnlijk [_{PP} op niemand]_i [_{AP} erg dol *t_i*] is. [negation]
 that Marie probably of nobody very fond is
 ‘that Marie isn’t very fond of anybody.’

The examples in (6) illustrate contrastive topic/focus movement; see, e.g. Neeleman & Vermeulen (2012). These examples show that the relative order of contrastive topics/foci and modal adverbs does not provide a reliable test for distinguishing topic/focus movement from A-scrambling, but they do evoke very distinct intonation patterns: while contrastive topics/foci require contrastive accent, phrases moved by A-scrambling are normally deaccented. Small capitals and double underlining indicate the A and B-accent used for, respectively, contrastive foci and topics (Neeleman & Van de Koot 2008); for our present goal, it suffices to say that examples involving contrastive accent should be eliminated from the study of A-scrambling.

- (6) a. dat Marie waarschijnlijk [_{ook} op *PETER*]_i [_{AP} erg dol *t_i*] is. [focus]
 that Marie probably also of Peter very fond is
 ‘that Marie is probably also very fond of PETER.’
 b. Ik weet niet wat Marie van Jan vindt, maar ik weet wel ...
 I know not what Marie of Jan considers, but I know AFF
 dat ze [_{op} *Peter*]_i waarschijnlijk [_{VP} [_{AP} erg dol *t_i*] is]]. [contrastive topic]
 that she of Peter probably very fond is
 ‘I don’t know how Marie feels about Jan but I do know she’s probably very fond of Peter.’

A second step in clarifying the picture is the realization that even a more restricted notion of A-scrambling still applies to a wider set of phenomena. This is clear from the fact -already illustrated in (3) and (4)- that though the leftward movement of subjects and objects is constrained in a similar way, they nevertheless target different landing sites: the subject moves into the regular subject position right-adjacent to C while the object moves into a position more to the right (close to the modal adverb). However, even if we restrict ourselves to objects, it is clear that not all leftward shifts are the same. The examples in (7) show that nominal objects can occupy various positions in the clause, that is, be moved over shorter or longer distances. Although this is often ignored in the literature, there are reasons to assume that the orders in (7b-c) are the result of different kinds of A-scrambling; see Schaeffer (2000: ch.2) and Broekhuis (2008: ch.2-3) for two different implementations of the same idea.⁵

- (7) a. dat ik waarschijnlijk nauwkeurig zijn boek moet lezen. [no A-scrambling]
 that I probably closely his book must read
 ‘that I probably must read his book closely.’
 b. dat ik waarschijnlijk zijn boek nauwkeurig moet lezen. [short A-scrambling]
 c. dat ik zijn boek waarschijnlijk nauwkeurig moet lezen. [canonical A-scrambling]

Short A-scrambling in (7b) just moves the object across a VP-adverb expressing manner, while canonical A-scrambling in (7c) also moves it across a clause adverb expressing modality; only in the latter case must the object belong to the presupposition of the clause. The earlier literature on object scrambling focused on canonical A-scrambling, i.e. on cases in which the object crosses a clause adverb that demarcates the boundary of the *vP* in the structures in (2). The

⁵ Schaeffer and Broekhuis differ in that the latter claims that (7a) and (7b) are both derived by short A-movement: the alternate word orders are the result of the order of merger of the VP-adverb *nauwkeurig* ‘closely’ and object movement: [_{ADV_{VP}} OBJ_i [_{VP} V *t_i*]] arises if object movement precedes merger of the adverb, and [OBJ_i ADV_{VP} [_{VP} V *t_i*]] if it follows it. As the focus of this paper is on canonical A-scrambling, I refer the reader to the original works for the details of the analyses.

terminology used in the non-generative literature was of course different, often more semantically based. Verhagen's (1986: ch.4) discussion, for instance, is phrased in discourse-related terms such as COMMENT, and the relevant adverbs are not characterized as clause adverbs but as COMMENT MODIFIERS. It is important to note that there is no one-to-one correspondence between the syntactic notion "vP" and the discourse-representational notion "comment", as is clear from the fact that vPs but not comments can contain presuppositional material: even pronouns can occur vP-internally when embedded within a PP-complement: cf. *dat zij waarschijnlijk [op hem wil wachten]* 'that she probably wants to wait for him'. Example (9d) below will show that there are also cases in which a presuppositional object NP remains within vP.

Short A-scrambling has not been extensively investigated so far, but it seems that it may serve various purposes. For instance, the examples in (8) differ in the relative scope of the object and the adverbial phrase: (8a) involves wide scope of the adverbial phrase, so that Jan is said to read 14 poems every week (or fewer if he rereads some of the poems); in (8b), the object has wide scope, so that Jan is said to read the same two poems over and over again.

- (8) a. dat Jan waarschijnlijk elke dag twee gedichten leest. [∀ > two poems]
 that Jan probably every day two poems reads
 'that Jan is probably reading two poems a day.'
 b. dat Jan waarschijnlijk twee gedichten elke dag leest. [two poems > ∀]

For our present purpose it is important to note that examples containing an object preceding a VP-adverbial like *nauwkeurig* 'closely' in (7) cannot be used for evaluating the claim that a canonically scrambled object must be part of the presupposition of the clause, as they can also be derived by short A-scrambling.

It is tempting to conclude from the preceding that canonical A-scrambling is obligatory when the object belongs to the presupposition of the clause, but (2e) has already illustrated that this position cannot be upheld: canonical A-scrambling of presuppositional definite NPs is normally blocked when it would change the unmarked order of the nominal arguments in the clause. This is illustrated again by the double-object constructions in (9). Examples (9b&c) show that the indirect and the direct object can both scramble to the left of the clause adverb *waarschijnlijk* (provided they are presuppositional). Example (9d) shows, however, that the direct object cannot scramble across the indirect object (not even when the former is presuppositional); see Haegeman (1993a) and Den Dikken (1995) for similar facts with pronominal direct objects. This means that (9a) is information-structurally ambiguous; the direct object need not be part of the focus of the clause but can also be part of its presupposition.

- (9) a. dat hij waarschijnlijk [zijn moeder het boek heeft gegeven].
 that he probably his mother the book has given
 'that he has probably given his mother the book.'
 b. dat hij zijn moeder_i waarschijnlijk [_{t_i} het boek heeft gegeven].
 c. dat hij zijn moeder_i het boek_j waarschijnlijk [_{t_i} _{t_j} heeft gegeven].
 d. *dat hij het boek_j waarschijnlijk [zijn moeder _{t_j} heeft gegeven].

Note that we are dealing with a language-specific constraint here, as German does allow canonical A-scrambling of the direct across the indirect object. On the hypothesis that canonical scrambling and Icelandic object shift involve the same A-movement operation, the constraint illustrated in (9) can be seen as part of a more comprehensive constraint known as Holmberg's generalization, which in its current formulation forbids A-movement across phonetic material within vP; cf. Holmberg (1999). This generalization was originally proposed to account for the fact that object shift cannot cross the non-finite verb in VO-languages like Icelandic. This is illustrated by (10a-b), taken from Holmberg (1986): while (10a) shows that object shift is possible provided the main verb is finite and moved into a vP-external position, (10b) shows

that it is blocked when the main verb is a participle and still located within the vP (indicated by square brackets). The Yiddish example in (10b'), taken from Diesing (1997), again illustrates the language-specificity of Holmberg's generalization; see Broekhuis (2020) for a more detailed discussion of the examples in (10).

- (10) a. Jón keypti <bókina> ekki [_v <bókina>]. [Icelandic]
 Jón bought the.book not
 b. Jón hefur <*bókina> ekki [keypt <bókina>].
 Jón has the.book not bought
 b'. Maks hot <dos bukh> nit [geleyent <dos bukh>]. [Yiddish]
 Maks has the book not read

It is not my intention to substantiate here the idea that canonical A-scrambling of objects and object shift should receive a uniform account in terms of A-movement, which I do believe to be the null hypothesis from a generative perspective. The cross-linguistic evidence discussed above still suggests, however, that any theory of leftward A-movement of objects should be able to account for specific cases in which presuppositional definite objects remain within vP, an option that more recent theories of object scrambling/shift do allow. The effect-on-output restriction in Chomsky (2001: 33-37), for instance, forces a presuppositional reading of objects that have been A-moved into their case position external to vP, but is silent on the interpretation of *in-situ* objects, and therefore (correctly) allows presuppositional readings for the definite object NP in (10b). OT-approaches, such as developed in Broekhuis (2000/2008) or Engels & Vikner (2014) can easily account for true optionality of object scrambling by appealing to the independently motivated notion of CONSTRAINT TIE, i.e. by claiming that the constraints favoring or disfavoring canonical scrambling are equally ranked. The impossibility of canonical scrambling of indefinite NPs and the obligatory canonical scrambling of definite pronouns illustrated in (3a,c) can be accounted for by appealing to constraints that do block/favor movement of these phrases but not of definite nominal objects.⁶ It is important to point this out, as De Hoop (2016) contended that the claim in Broekhuis (2008) and Broekhuis & Corver (2020) that presuppositional definite objects normally undergo A-scrambling in Dutch is not empirically motivated but exposes a theoretical bias. The conclusion above that any theory of A-scrambling should allow presuppositional definite objects to remain inside vP in certain cases underlines that De Hoop's contention is incorrect: Broekhuis & Corver simply share the judgments on the relevant examples found in the earlier literature, which ultimately led to the formulation of this generalization in Verhagen (1986).

3 Van der Does & De Hoop (1998)

Van der Does & De Hoop follow the earlier literature on object scrambling in adopting the distinctions illustrated in (3): nonspecific indefinites are not allowed to scramble, definite pronouns must scramble, while definite NPs scramble “optionally”. They diverge from this literature, however, in taking the notion OPTIONAL literally: “definites may either occupy the

⁶ Such an additional constraint is needed for definite object pronouns anyway, as there is reason for assuming that they occupy a higher position in the clause than definite object NPs: object pronouns normally occur right-adjacent to the subject (or the finite verb in non-inverted main clauses) and consequently precede all higher clause adverbs in the middle field of the clause, while definite object NPs can easily follow such adverbs: *dat Jan <het> volgens Marie <*het> zeker zal kopen* ‘that, according to Marie, Jan will certainly buy it’. It seems that canonical scrambling of definite pronouns is obligatorily followed by an additional movement into the position immediately following the canonical subject position. Broekhuis (2008: 176-7) claims that this additional movement is of the A'-type, as it also applies to non-nominal pro-forms like *er/daar* ‘there’ and *toen* ‘then’; cf. Haider et al. (1995). See Broekhuis & Corver (2016: §13.4) for more discussion of pronoun scrambling.

scrambled or the unscrambled position relative to an adverb, and there does not seem to be a property of either the definite itself or the context in general that forces or prohibits scrambling” (p.399). The first claim that there are no properties of the definite itself that force or prohibit scrambling rests on cases like the ones in (11) with so-called weak definite NPs. Weak definites are non-referential definite NPs that form a collocation with the verb; the collocation functions as a complex predicate referring to frequently occurring events involving humans, like domestic activities and traveling; see Aguilar-Guevara (2014) and references cited there for extensive discussion.

- (11) a. dat ik <de was> nog <de was> moet doen.
 that I the laundry still must do
 ‘that I still have to do the laundry.’
 b. dat ik <de bus> altijd <de bus> neem.
 that I the bus always take.
 ‘that I always take the bus.’

Since the weak definite objects are not referential, Van der Does & De Hoop claim that theories of scrambling which require scrambled objects to be part of the discourse domain cannot account for the word-order variation found in (11). They seem to be jumping to conclusions, however, as the examples in (11) are in fact ambiguous; they can also be construed with a referential interpretation of the object, as is clear from the fact that the objects can easily be modified: *de was van gisteren* ‘yesterday’s laundry’ and *de eerste bus* ‘the first bus’. The question is not whether the two orders are acceptable, but whether they are both acceptable on the intended weak reading of the definite object. According to my own judgments, this reading is much better when the definite object is not scrambled, which is in line with the outcome of the rating experiment reported in De Swart & Van Bergen (2014): “Sentences with scrambled weak definites were rated significantly lower in comparison to all other conditions” (p.292). That scrambled weak definites are less favored is also shown by a Google search [01\24\2021] on the two strings [*elke dag de was doen*] and [*de was elke dag doen*], in which the frequency adverbial *elke dag* ‘every day’ favors the weak reading; the unscrambled order resulted in approximately 60 unique hits while there were no cases with the scrambled order at all; scrambling across *elke dag* is also exceptional for the collocations *de krant lezen* ‘to read the newspaper’, *de ramen lappen* ‘to clean the windows’ and *de bus/trein nemen* ‘to take the bus/train’. I therefore conclude that the claim that scrambling of weak definites is truly optional is too strong: it is exceptional, if possible at all.

The second claim that there are no properties of the context forcing or prohibiting scrambling is motivated by cases such as (12), for which Van der Does & De Hoop claim that the scrambled and the unscrambled version are synonymous. It is not *a priori* clear, however, whether the presumed synonymy is relevant for evaluating the claim that contextual properties are involved in the distribution of definite objects, as Van der Does & De Hoop seem to think about meaning strictly in terms of type logic, i.e. they do not take information structure into account. However, even if we assume for the sake of the argument that the authors did have contextual factors in mind, there are various problems with using (12) as an argument against the claim that scrambling depends on contextual factors.

- (12) dat ik <de kraker> gisteren <de kraker> heb gesproken.
 that I the squatter yesterday have spoken
 ‘that I talked to the squatter yesterday.’

Let us start with two problems. First, it seems to me that the preferred readings of the two examples differ in the way predicted by the traditional view: the scrambled object is very likely part of the domain of discourse. This is also reflected by the findings in the corpus study

reported in Van Bergen & De Swart (2009): “anaphoric and non-anaphoric objects both have a preference for the unscrambled position, although anaphoric objects do scramble more often than non-anaphoric objects (22% versus 7% in scrambled position, respectively; Fisher exact test, $p < .0001$).”

That there is a (significant) effect of anaphoricity does of course not take the edge off the fact that there are many cases in which an anaphoric definite object is not scrambled. This leads us to the second, more serious, problem with examples such as (12): they can simply not be used for testing claims that scrambling depends on contextual factors because they do not contain the right adverb type. Van der Does & De Hoop (1998) –and the same holds for the corpus/experimental studies to be discussed later– use SCRAMBLING as “a descriptive term for the occurrence of an object to the left of an adverb in an SOV-language” (p.393). Section 2 has already discussed that this means that scrambling is used as a cover term for at least two distinct kinds of A-scrambling which we referred to as short scrambling, which may cross a VP-adverb but not a clause adverb, and canonical scrambling, which also crosses the clause adverbs located at the left border of vP. It is only canonical scrambling that has the property that it cannot apply to non-presuppositional definite objects. Therefore, even if the two orders in (12) were informational-structurally equivalent, this would not refute the claim that canonical scrambling depends on the information structure of the clause, as we first have to show that the time adverb *gisteren* ‘yesterday’ is of the right kind.

Now consider the examples in (13). Example (13a) shows that time adverbs may either precede or follow the modal adverbs: they function as clause adverbs in the former and as a VP-adverbs in the latter case. The main restriction on the co-occurrence of the two kinds is that the time interval/point referred to by the VP-adverb is included in the time interval referred to by the clause adverbial; cf. Broekhuis & Corver (2016: §8.2.3). Since this condition is trivially satisfied when the clause contains a single time adverbial, as in (13b), the adverbial can then occur in either position.

- (13) a. Jan is *gisteren* waarschijnlijk om drie uur vertrokken.
 Jan is yesterday probably at 3 o'clock left
 ‘Jan probably left at 3 o'clock yesterday.’
 b. Jan is <*gisteren*> waarschijnlijk <*gisteren*> vertrokken.
 Jan is yesterday probably left
 ‘Jan probably left yesterday.’

The cases in (14) provide instances with *morgen* ‘tomorrow’ used as a clause adverb: (14b) shows that canonical scrambling can place the definite object in between the time and the modal adverb. The interpretations of these examples seem to be in line with the predictions made by the more traditional literature; in (14a), the definite object is part of (\subset) of the focus and in (14b) of the presupposition of the clause. The crucial point for what follows is that the modal adverb is not mandatory: its omission leads to the linear string in (14c), which is ambiguous between the two information-structural readings in (14a&b).⁷

⁷ Since canonical scrambling leads to deaccentuation of the object, example (14c) can be disambiguated by means of intonation; see Verhagen (1986) and Neeleman & Reinhart (1998) for more discussion. As Van der Does & De Hoop (1998) and the corpus/experimental studies to be reviewed below have nothing to say about intonation, I will not discuss this here.

- (14) a. dat ik morgen waarschijnlijk de kraker spreek. [de kraker \subset focus]
 that I tomorrow probably the squatter speak.to
 b. dat ik morgen de kraker waarschijnlijk spreek. [de kraker \subset presupposition]
 that I tomorrow the squatter probably speak.to
 c. dat ik morgen de kraker spreek. [ambiguous]
 that I tomorrow the squatter speak.to
 ‘that I will (probably) speak to the squatter tomorrow.’

A similar ambiguity arises when the object precedes a temporal VP-adverb; (15a&b) show that a definite object can precede the VP-adverb on both readings. Consequently, omission of the modal adverb, as in (15c), leads to ambiguity.

- (15) a. dat ik waarschijnlijk de kraker morgen spreek. [de kraker \subset focus]
 that I probably the squatter tomorrow speak.to
 b. dat ik de kraker waarschijnlijk morgen spreek. [de kraker \subset presupposition]
 that I the squatter probably tomorrow speak.to
 c. dat ik de kraker morgen spreek. [ambiguous]
 that I the squatter tomorrow speak.to
 ‘that I will (probably) speak to the squatter tomorrow.’

If we now compare the (c)-examples in (14) and (15), we see that the same ambiguity arises with temporal clause and VP-adverbs. So my divergence of view with Van der Does & De Hoop (1998) therefore does not concern their interpretation of example (12) but the fact that they use this example as an argument against the traditional claim that canonical scrambling depends on contextual factors. They fail to select the right kind of clause adverbial: temporal VP-adverbials are internal to ν P and can therefore be crossed by short A-scrambling while temporal clause adverbials are not sufficiently close to the left ν P-boundary to be crossed by canonical A-scrambling; see also Broekhuis & Corver (2016: 308).⁸

4 Alternatives to introspection research

De Hoop (2000/2003) claims again that scrambling is “truly” optional: she presents a probabilistic OT-model which predicts that 66% of all anaphoric and 50% of all non-anaphoric definite NPs scramble in actual speech (without any difference in meaning). Although the articles cite a number of examples from a children’s book, the claim is essentially based on De Hoop’s own intuitions. The two articles include a number of examples with clause adverbs (e.g. *toch* ‘nevertheless’ and *helaas* ‘unfortunately’), which shows that De Hoop believes scrambling of definite object NPs to be “truly” optional in such cases as well. This means that the discussion in Section 3 may perhaps have lessened the presumed disagreement on the data but this is not sufficient to reject the claim that scrambling of definite object NPs is “truly” optional. This raises the question as to how to settle the remaining conflicting judgments.

One way would be to rely on intersubjectivity: as most researchers have committed themselves to the claim that canonical A-scrambling in Dutch is discourse related, dissident opinions such as De Hoop’s can be put aside as incorrect or as applying to a non-standard variety of Dutch. Such a “brute force” solution is of course undesirable, and should only be invoked if all other options fail. Another way would be to declare the relevant data to be “unclear” and settle their grammaticality status by “letting the theory decide”: cf. Newmeyer (1983: §2.2). This option is only available, however, if researchers share sufficient common ground, which is unlikely to hold in this case as researchers involved in the study of scrambling

⁸ Definite object pronouns must precede the temporal clause adverbs: *dat ik <hem> morgen <*hem> waarschijnlijk spreek* ‘that I will probably speak to him tomorrow’. This supports the suggestion in note 6 that they are subject to an additional movement into the position right-adjacent to the canonical subject position.

have quite different theoretical backgrounds. In fact, this option may not even be available for individual researchers, as most present-day theories include specific parameters aiming at accounting for the cross-linguistic variation found in Germanic scrambling and object shift; the proposals in Chomsky (2001) and Broekhuis (2008) discussed in Section 2, for instance, allow a wide variety of language systems and therefore do not provide much guidance in declaring certain viable output forms (un)grammatical for a specific language. This holds to an even greater extent in the case of De Hoop's (2000/2003) probabilistic approach, which only says something about the *likelihood* (0-100%) that a specific object NP is scrambled in a certain language. Thus, if we want to avoid the "brute force" solution, we have no choice but to examine the judgments made in the earlier literature by comparing them with the results of corpus research and/or research in an experimental setting. We will see, however, that the attempts undertaken so far are not without their problems either. Section 4.1 reported a corpus study, after which 4.2 to 4.4 continue with a number of experimental studies.

4.1 Corpus research

Van Bergen & De Swart (2009/2010) investigate scrambling of direct objects in the speech of Dutch speakers from the Netherlands. This is done by means of the syntactically annotated part of the corpus of spoken Dutch (CGN), which includes types of speech ranging from very informal to very formal. Van Bergen & De Swart (2009) aim at determining to what extent definiteness and anaphoricity play a role in order to evaluate earlier scrambling theories; there is, however, no mention of the role of information structure in terms of the presupposition-focus division of the clause, which was taken to be the decisive factor in the more traditional approaches to object scrambling such as Verhagen (1986).

This section mainly focuses on the corpus study reported in Van Bergen & De Swart (2009).⁹ The sample extracted from the CGN-corpus consists of transitive constructions with a nominal direct object directly preceded or followed by an adverb, produced by speakers from the Netherlands. Adverbial PPs and pro-forms like *er/daar* 'there' are not included in the sample, probably because they exhibit special placement restrictions; the same holds for adverbial clauses although this is not explicitly mentioned. As there are no further restrictions on the kind of adverb, we may conclude that their study adopts the wide definition of scrambling also used in Van der Does & De Hoop (1998); it consequently runs up against the same objections, more specifically that examples such as (12) above cannot be used for testing claims that canonical A-scrambling depends on contextual factors, as there is no guarantee that the selected sentences contain the right adverb type. A conspicuous restriction on the sample is that sentences with an object in between adverbs are "left out as they could not be uniquely classified as scrambled or unscrambled" (Van Bergen & De Swart 2010: 270). The desirability of this decision is highly debatable since such cases would count as "scrambled" in most theories. It cannot be determined, however, whether this restriction has had a significant effect on the outcome of their research, as the number of cases involved is not mentioned.

In the resulting subcorpus, the object-adverb order and adverb-object order occur in, respectively, 53% and 47% of the cases. A smaller sample of 2900 utterances with the same ratio of the two word orders was created for further manual annotation and analysis. The direct objects in the sample are divided into four main groups: pronominals (definite and reflexive/reciprocal pronouns), proper nouns, definite NPs, and indefinite NPs. The pronominals ($n = 1287$; 44.4%) and indefinite NPs ($n = 1187$; 40.1%) clearly outweigh the two

⁹ Van Bergen & De Swart (2010) confirms the findings in Van Bergen and De Swart (2009) on the basis of a larger sample. It also investigates whether the effect of definiteness on scrambling can be reduced to the effect of weight (in terms of the number of characters) in a logistic regression model, but concludes that the degree of definiteness is a better predictor than object length. In addition, they found that scrambling of proper nouns is affected by animacy, length and stress.

remaining groups ($n=426$) and exhibit strict word order. The pronouns precede the adverb in 99% of the cases, which concurs with my earlier observation that they occupy a high position in the clause, right-adjacent to the canonical subject position: cf. notes 6 and 8. The indefinite NPs occur to the right of the adverbs in 98% of the cases, which is consistent with the generally accepted claim that they do not scramble. We will therefore put the utterances with pronouns and indefinite NPs aside and concentrate on utterances with definite NPs.¹⁰

The sample contains 367 utterances with a definite NP (12.6%). A division of the object NPs based on their determiner (article, demonstrative/possessive pronoun) or numeral quantifier (e.g. *alle* ‘every’) does not reveal great differences in word order, although it should be mentioned that definite NPs with an article scramble significantly less than those with a demonstrative (8% versus 19%); the definite NPs precede the adverb as a group in only 12% of the cases. Another division is made between anaphoric (114 utterances; 3.9%) and non-anaphoric (253 utterances; 0.087%) NPs; anaphoricity was established manually by checking whether the NP was mentioned in the preceding discourse. These two groups do exhibit a significant difference in word order: 22% of the anaphoric NPs precede an adverb while this is the case with only 7% of the non-anaphoric definite NPs. Van Bergen & De Swart (2009: §.3.3) further note that the overall number of anaphoric definite object NPs “is comparatively low, something which can be explained by the fact that in those cases speakers are more likely to use pronouns instead; cf. the high overall number of pronominal objects”. A third division involves only the group of definite NPs with an article, which were divided in referential and non-referential expressions; the latter group probably involves the weak definites discussed earlier. The manually enriched sample is not publicly available, which is unfortunate as this makes a further analysis of the utterances impossible: we will see below, for instance, that to investigate whether the various adverb types in the sample have a significant effect on word order would have been welcome.

We now turn to the theoretical implications of the findings summarized above. Van Bergen & De Swart (2009) first deal with the prediction made in Diesing (1992/1997) and, especially, Diesing & Jelinek (1995) that only non-existential (\approx non-specific indefinite) object NPs can be in a ν P-internal position (that is, under existential closure). They conclude that this prediction is refuted by the fact that definite NPs are in an unscrambled position in 88% of the cases. Two objections can be raised against this conclusion. First, a minor point: Van Bergen and De Swart mention that the set of definite NPs includes about 87 non-referential (weak) NPs but these are simply irrelevant to the present discussion; cf. Diesing & Jelinek (1995: 130-1), Van der Does & De Hoop (1998: 398), and also Van Bergen & De Swart (2010:269). When we exclude these cases, the relevant portion of unscrambled definite NPs will be slightly lower (84.7%). A more important point is that Diesing & Jelinek’s prediction that definite NPs must be ν P-external does not entail that they must cross an adverb (if present): if the adverb is a clause adverb higher in the structure than the ν P-external landing site of the object, no crossing is expected. The point against Diesing & Jelinek’s (1995) prediction can therefore only be made if it can be shown that (referential) definite object NPs can follow VP-adverbs or lower clause adverbs like the modal *waarschijnlijk* ‘probably’. As it seems uncontroversial that this is possible in Dutch (cf. (1b)), I agree with the conclusion that the prediction cannot be upheld although the argumentation given clearly does not come up the mark.

¹⁰ There were also 59 utterances with a proper noun (2%). I will not discuss the distribution of proper nouns here but it should be pointed out that they were found in the object-adverb order much more often than definite NPs (53% versus 12%). Although I do not take a firm stand on the issue, this might follow from theories that assign proper nouns a similar structure as definite pronouns, in the sense that they both undergo N-to-D movement; see Longobardi (1994) and much later work.

What are the implications for the hypothesis that anaphoric definite object NPs must scramble? The proportion of scrambling is 22 % in this case. Exclusion of the non-referential noun phrases might again raise this number slightly, although this is difficult to establish. We may therefore conclude that the corpus data refute the naive assumption that anaphoric definite object NPs must cross an adverb (if present). However, one might want to save the hypothesis by adopting the more sophisticated view that anaphoric definite object NPs just cross the VP-adverbs and the lower clause adverbs. For reasons to be discussed shortly, I believe that this would still leave a residue of unscrambled anaphoric definite object NPs, so that the conclusion that the (revised) hypothesis cannot be maintained is probably correct. Bergen & De Swart's argumentation leading to this conclusion is, however, not without flaw.

The more traditional theories maintaining that (canonical) scrambling is determined by the information structure of the clause are not discussed in Van Bergen & De Swart (2009). The implementations of this idea in the more recent versions of generative grammar resemble Diesing & Jelinek's (1995) proposal in that presuppositional definite object NPs are assumed to move out of the *vP* into their case positions. An important difference with Diesing & Jelinek's proposal is, however, that this is not a hard-and-fast rule: Section 2 has already argued that object shift (including scrambling) is a language-specific phenomenon subject to language-specific conditions. We therefore expect definite object NPs to remain within the *vP* in specific languages or, in languages that do have object shift, under specific circumstances. Here, we focus on the constraint related to information structure in terms of the presupposition-focus division in the clause, but there are more constraints such as Holmberg's generalization related to order preservation. The notion of presuppositionality bears, of course, a certain resemblance to the notion of anaphoricity but is more sophisticated: it is not the mere mention of an NP in the earlier discourse but its reference to an "activated" discourse entity which determines its presuppositionality; the (new information) focus, on the other hand, is the speaker's own contribution to the information already available in the ongoing discourse. The difference between anaphoricity and presuppositionality can be nicely illustrated by appealing to the earlier mentioned fact that the number of anaphoric definite object NPs is relatively low, which "can be explained by the fact that in those cases speakers are more likely to use pronouns instead" (Van Bergen & De Swart 2009. §3.3). This suggests that anaphoric definite NPs are mainly used when this is needed to avoid lack of clarity, for instance, to *reactivate* some discourse entity that has not been mentioned for a while. If so, anaphoric definite NPs are normally not part of the presupposition but of the focus of the clause, and are therefore expected not to scramble very often, which confirms what Van Bergen & De Swart (2009/2010) found.

It should be clear that the results of the present corpus study do not allow an evaluation of scrambling theories in general, notwithstanding that the refutation of the theories based on the notions of definiteness and anaphoricity is in fact justified. The main reason is that such evaluations require a subcategorization of adverbs involving at least the following three categories which are of a different order for establishing scrambling to a position external to *vP*: high clause-adverbs (which are irrelevant), low clause-adverbs at the border of *vP* (which are conclusive) and VP-adverbs (which are indicative but not decisive). A complicating factor is that certain adverbials are flexible in their use: time (and place) adverbials, for instance, can either function as high clause adverbs or as VP-adverbs, and often their function cannot be determined purely by their position in the clause, leading to true structural ambiguity; see the discussion of the examples in (13)-(15). The present study did not address the more traditional theories of scrambling based on the information structure of the clause: although we have just seen that it may in fact weakly support such theories, it is not adequately equipped for a conclusive evaluation of them, as it presupposes a careful information-structural analysis of the utterances in the extracted sample in discourse. This can only be remedied by performing an

information-structural analysis of a supplementary sample containing transitive utterances with a small set of low clause adverbs such as the modal *waarschijnlijk* ‘probably’ in context.

4.2 Experimental approaches I: definite object NPs

In an unpublished follow-up study, De Swart & Van Bergen (2011) purport to provide experimental support for their earlier findings in the corpus data discussed in Section 4.1. They carried out one rating test, and two production tests. The results show that, although the rating test revealed that sentences with scrambled and unscrambled definite object NPs are rated equally, there is a strong preference for the unscrambled order in production. They construe this as support for their earlier conclusion that approaches to scrambling based on definiteness (Diesing & Jelinek 1995) or anaphoricity (Neeleman & Reinhardt 1998) cannot be upheld. They again fail to discuss approaches based on the presupposition-focus division of the clause along the lines in Verhagen (1986), Schaeffer (2000), or Broekhuis (2000/2008).

The experimental studies differ from the earlier corpus studies in that they allow us to determine the appropriateness of the data for theory evaluation, as the input of the tests is given in the appendices. What is especially problematical here is that the three experiments target time adverbials only, as we have seen that such adverbs cannot only be construed as VP-adverbs but also as high clause adverbs; see the discussion of (13)–(15) in Section 3. This makes the experiments unsuitable for testing theories addressing object movement into a *vP*-external position, for two reasons: (i) objects preceding a temporal VP-adverb may have moved into either a *vP*-internal position (short A-scrambling) or a *vP*-external position (canonical A-scrambling); (ii) objects following a temporal clause adverb may likewise have moved into either a *vP*-internal or a *vP*-external position (or remain in situ). The required distinction between short and canonical scrambling can only be made by considering the order of the object and a low clause adverb such as *waarschijnlijk* ‘probably’.

In this light, it need not surprise us that **the rating test** shows that subjects rate the two word orders as equally plausible in the case of (referential) definite object NPs, as the two orders can be analyzed in various ways. Sentences with the ADV<NP order are four-ways structurally ambiguous: they can be analyzed as involving a temporal clause adverb with canonical, short or no A-scrambling of the object, or as involving a temporal VP-adverb with the object *in-situ*. Sentences with the NP<ADV order are at least two-ways ambiguous: they can be analyzed as involving a temporal VP-adverb with either long or short A-scrambling. As all structures are equally grammatical, we may expect that, out of context, no significant differences in rating will be found. Note in passing that sentences with a weak (non-referential) object NP are rated significantly lower in the NP<ADV than in the ADV<NP order. This can be accounted for if we adopt the standard assumption that weak object NPs stay within *vP*. The fact the NP<ADV order is still rated relatively high can mean various things. The most promising approach in my view is that, as already suggested in the discussion of (11) in Section 3, the test sentences are actually ambiguous in that the presumed weak NP can also be construed as strong (referential), although one might also entertain the idea that weak object NPs at least marginally allow short A-scrambling.

The first production test is a **sentence-completion task** based on the same sentences as used in the rating test. Subject were presented with an incomplete sentence on a computer screen such as *Jan zegt dat ...* ‘Jan says that ...’, which they had to complete by using a list of elements presented later consisting of (i) a nominative pronoun (e.g. *hij* ‘he’), (ii) a time adverb (e.g., *gisteren* ‘yesterday’), (iii) a definite noun phrase (e.g., *het boek* ‘the book’), and (iv) an infinitival verb (e.g., *lezen* ‘to read’). These elements were presented in eight different orders, in which the adverb and the definite NP never followed each other directly. The answers were given orally and were recorded. Utterances were only coded as (UN)SCRAMBLED when the

adverb and the NP were adjacent in the output. Utterances that did not fit the template (e.g. with verbs in non-final-position) or were not coded for (UN)SCRAMBLED were excluded from the analysis reported in the paper. The results show a significant preference for the unscrambled order (regardless of the order of the adverb and the NP in the input list). The authors claim that this makes it unlikely that definiteness is the determining factor in the production of the two orders. However, on the assumption that the scrambled object targets a position to the right of the position of the high sentence adverbs makes this conclusion inconclusive: speakers might also have a preference for inserting time adverbs as high clause adverbs in out-of-the blue contexts, as temporal clause adverbs are used to temporally locate the proposition expressed by the sentence in time; see Broekhuis & Corver (2016: §8.2.3) for discussion.

It stands to reason that the results of the first production test are unfit for evaluating the role of anaphoricity since no context is provided. The second production test therefore investigated the word-order variation of adverbs and definite object NPs in the context of a three-sentence story. The second sentence in the stories varies in that it either introduces or does not introduce a referent for the definite object NP in the third (target) sentence. The stories were pre-tested by means of a rating study on equal plausibility and coherence. The actual test was a **sentence-completion task** along the same lines as the first one. The results again show a significant preference for the unscrambled order, regardless of whether the second sentence introduces the referent of the definite object NP. The authors take this to contradict the claim that anaphoric definite object NPs must scramble. Several objections can be made to the test in its present form. The authors themselves mention that even if the second sentence does not introduce the referent of the definite NP by means of an indefinite NP, this referent can still be implied by the context it sets up, which may be relevant for the fact that the anaphoric and non-anaphoric definite object NPs did not exhibit different behavior with respect to word order. A second objection is that the target sentences with an anaphoric definite object NP are rather unnatural in that in natural speech the use of a pronoun would be the more normal option. But the most important point is, of course, that the results only refute the naive version of the criticized approach: if we assume that A-scrambling of the object targets a *vP*-external position to the right of the position of the high sentence adverbs, the preference for the ADV<NP order can be accounted for in a similar fashion suggested earlier: there might be a preference for using the time adverb as a clause adverb rather than as a VP-adverb.

Our discussion of the experiments reported in De Swart & Van Bergen (2011) has shown that the reliance on temporal adverbs disqualifies the results for evaluating approaches to object scrambling built on the assumption that this movement targets a designated position external to *vP*. The reason for this is that time adverbs can be used either as clause adverbs, in which case they precede the landing site of canonically scrambled objects, or as VP-adverbs, in which case they follow this position. Note that this problem for the experiments can be overcome in principle when the temporal adverbs in the tests are replaced by lower clause adverbs.

4.3 Experimental approaches II: adverb types

Schoenmakers & De Swart (2019) again discuss scrambling of definite object NPs but now the focus is not on properties of the objects themselves: they simply adopt Van der Does & De Hoop's (1998) claim "that scrambling is truly optional for definite objects". The question Schoenmakers & De Swart want to answer is instead whether there are object-external factors prompting speakers to choose between the alternate word orders. The article is prompting for our present discussion, as it focuses on the distinction between clause and VP-adverbs, which are referred to as "high" and "low" adverbs.

Scrambling is construed in the same naive way as before: the simple inversion of the order of an object and an adverb. The hypothesis that object scrambling is truly optional would

therefore presumably lead to the expectation that there is no difference in the frequency of object scrambling across clause or VP-adverbs. Schoenmakers & De Swart (2019:126) sidestep this by adopting the economy hypothesis that “movement of an object across a high adverb is more costly than movement across a low one, because the distance between the object’s base and target position is larger” or, more precisely, because it would involve more derivational steps; see Schaeffer (2000: ch.2) and Broekhuis (2008: ch.2-3) for two different implementations of this idea. This is advantageous as Schoenmakers & De Swart now seem to adopt the distinction between short and canonical A-scrambling established on the basis of the examples in (7) in Section 2; this potentially removes one of the main deficiencies of the earlier experimental studies.¹¹

Schoenmakers & De Swart (2019:126) claim that the economy hypothesis leads to the following two predictions for their experiments: “(i) definite objects are better in the scrambled position in sentences with a low adverb than in sentences with a high adverb [in their rating experiment IA], and (ii) definite objects appear in the scrambled position more often in sentences with a low adverb than in sentences with a high adverb [in their production experiment IB]”. The rating and production experiments were conducted in the same way as those in De Swart & Van Bergen (2011). The stimuli of the rating test consisted of 28 transitive sentences with a definite object NP and a high or a low adverb: half the stimuli for the rating test were adapted from the earlier study, while the other half (probably those with a low adverb) were newly formulated from scratch. The production test was a sentence-completion based on 24 sentences used for the rating test.

The results of the **rating test** shows that the participants rated all sentences high. The average is approximately 6 on a 7-point scale in the following order (high to low): (i) NP < low adverb; (ii) NP < high adverb; (iii) high adverb < NP; (iv) low adverb < NP. Schoenmakers & De Swart did not find any significant effect for adverb type: sentences with and without scrambling were rated more or less the same for each adverb type. There was, however, a significant effect for the interaction of word order and adverb type: NP < adverb was preferred in the case of low (but not high) adverbs. This would suggest that the economy hypothesis is not really supported: there is some effect of the length of the path (or number of steps) involved in object scrambling but it is only a marginal one. The authors conclude that the results again support the earlier findings that Dutch native speakers happily accept both word orders. Although I am more than willing to accept this conclusion, I believe that at least two general remarks are in order here.

The first remark pertains to the premise that economy considerations affect the rating of the stimuli. Economy arguments were originally introduced to account for the fact that sentences involving fewer/shorter derivational steps are preferred to *competing* sentences involving more/longer derivational steps.¹² There is, however, no evidence suggesting that sentences resulting from a more economical derivation are preferred because they are also easier to process. Even if this was the case, such complexity might be irrelevant in the case of scrambling if word order has a specific discourse function, as we would then no longer be dealing with *competing* sentences in terms of usage conditions. In short, I see no *a priori* reason to expect differences in the rating of sentences with canonical or short scrambling.

The second remark concerns the presumed “high” adverbs used in the stimuli. Schoenmakers & De Swart mention that half of the stimuli (probably those with a “high”

¹¹ I will ignore the possibility briefly considered but not taken up by Schoenmakers & De Swart (2019:129) that direct objects are base-generated to the left of VP-adverbs.

¹² I put aside the fact that Chomsky (1995) claims that less economic derivations are in fact *blocked* by more economical ones, which implies that economy considerations of the sort under discussion cannot play a role in the processing of actual speech by definition.

adverb) were adapted from De Swart & Van Bergen (2011), while the other half were created from scratch. Although this suggests that they used temporal adverbs as representatives of the “high” adverbs, Gert-Jan Schoenmakers (p.c.) has informed me that the stimuli involved the true clause adverbs *blijkbaar* ‘apparently’, *gelukkig/helaas* ‘(un)fortunately’, *inderdaad* ‘indeed’. This is an improvement since it partly avoids the problem with respect to structural ambiguity arising from the use of temporal adverbs. Unfortunately, the selection of clause adverbs does not completely avoid structural ambiguity, as *gelukkig* ‘fortunately’ and *helaas* ‘unfortunately’ are high clause adverbs; cf. Barbiers (2018: 62). The word order ADV>NP is therefore three-ways ambiguous in the case of these adverbs: it may involve canonical, short or no A-scrambling. This means that some of the cases analyzed as “unscrambled” may in fact involve canonical scrambling. Note that this does not damage the conclusion drawn by the authors; it simply adds to their conclusion that the economy hypothesis is not supported by the outcome of the experiment.

The results of the **sentence-completion task** show that there is a clear effect of adverb type: definite object NPs preceded the “low” adverbs but followed the presumed “high” adverbs in the majority of the cases.¹³ The authors take this as support for the economy hypothesis but, again, there is reason to reject this conclusion. First, A-scrambling of the object across the “low” adverbs does not provide information about the actual position of the object, that is, it may result from either short or canonical A-scrambling. Second, the fact that an object follows a presumed “high” adverb does not imply that canonical A-scrambling did not apply, as (14b) has shown that canonically scrambled objects can follow temporal clause adverbs.

I therefore conclude that the experiments described in Schoenmakers & De Swart (2019) are not suitable for establishing the acceptability judgments of hearers on canonical A-scrambling nor for establishing the frequency of canonical A-scrambling of definite object NPs in production; the results can only be used for evaluating statements pertaining to linear order, not for statements pertaining to syntactic structure. The fact that the sentences were presented out of context would have made it difficult to use the result for evaluating scrambling theories appealing to anaphoricity or information structure (presupposition-focus division).

4.4 Experimental approaches III: negation

Schoenmakers & De Swart (2019) also intend to investigate the role of scope in scrambling of definite object NPs, an issue which is also taken up in Schoenmakers (2020). They conduct rating and sentence-completion tests of the by now familiar kind with the negative adverb *niet*, and find that “definite objects are scrambled across the negation adverb in the vast majority of utterances”. I will not discuss the results of their experiments at length, as I believe that they are at best only indirectly relevant to the discussion of canonical A-scrambling. This section provides my reasons for claiming this.

The experiments reported in Schoenmakers & De Swart (2019) are aligned with earlier studies on first (and second) language acquisition of object scrambling in Schaeffer (2000) and Unsworth (2005). The experiments in these studies involve scrambling across the negative element *niet* in examples of the type in (16). The assumption in all studies mentioned is that (16a) differs from (16b) only in the application of object scrambling, which is not independently substantiated and in fact contestable, as will become clear from my brief discussion below of constructions of the kind in (16a); see also Broekhuis & Den Dikken (2012: §8.1.3.4) and Broekhuis & Corver (2016: 1646ff.) for more discussion.

¹³ There was also an effect of the order of the input: the NP<ADV order occurred more frequently when the object preceded the adverb in the input than when it followed it, regardless the type of adverb. This is irrelevant for the point made here.

- (16) a. Jan heeft niet het boek gelezen. [constituent negation]
 Jan has not the book read
 ‘It was not the book that Jan read.’
 b. Jan heeft het boek niet gelezen. [sentence negation]
 Jan has the book not read
 ‘Jan hasn’t read the book.’

Examples such as (16a) cannot be interpreted as negative sentences; they are unacceptable with a neutral intonation pattern and marked as ungrammatical in Kraak (1966: 121). However, they do allow an interpretation as a positive sentence with constituent negation, provided the negated constituent is contrastively accented. Sentences of this kind are used especially when the speaker intends to correct a misconception of the addressee; cf. Haeseryn et al. (1997). Since the contrastive accent is not only restricted to the object NP but is also placed on the VP, the correction may involve either the object or the full VP; contrastive accent is indicated by small caps and the constituent in the scope of *niet* is underlined.

- (17) a. Jan heeft niet het BOEK gelezen maar de KRANT.
 Jan has not the book read but the newspaper
 ‘Jan has not read the book but the newspaper’
 b. Jan heeft niet het BOEK gelezen maar GESLAPEN.
 Jan has not the book read but slept
 ‘Jan has not read the book but the newspaper’

Dik (1997; §13.4) suggests that the constituent negator *niet* should be seen as a focus particle on a par with elements like *ook* ‘also’ and *alleen* ‘only’; these elements are all involved in counter-presuppositional focus constructions. Such constructions aim at modifying a subset of propositions which the speaker presupposes to be considered true by the addressee: correcting counter-presuppositional focus replaces a proposition held by the addressee containing X by a proposition containing Y instead of X; expanding counter-presuppositional focus adds Y to the proposition, while restricting counter-presuppositional focus removes Y; selecting counter-presuppositional focus makes a choice between X and Y. This is indicated in Table 1; (P_A)_s stands for the set of propositions supposedly held by the addressee and P_s stands for the revised set of propositions if the speaker succeeds in modifying (P_A)_s.

Table 1: Types of counter-presuppositional focus

	ORIGINAL SET (P _A) _s	MODIFIED SET P _s	EXPRESSION TYPE
CORRECTING	X	Y	not X, but Y!
EXPANDING	X	X and Y	also Y!
RESTRICTING	X and Y	X	only X!
SELECTING	X or Y	X	X!

Independent support for the claim that constituent negation should be distinguished from sentence negation is that they can co-occur within a single clause without canceling each other out (i.e. without resulting in a positive sentence); this is illustrated by the following examples, taken from Haeseryn et al. (1997: 1641).

- (18) a. Hij heeft de deur niet met OPZET niet dicht gedaan heeft.
 he has the door not on purpose not close done has
 ‘It was not on purpose that he didn’t close the door.’
 b. Ze is niet zonder reden niet komen opdagen.
 she is not without reason not come show.up
 ‘It was not without reason that she did not come up.’

Dik's claim that the constituent negator belongs to same group as the focus particles *ook* 'also' and *alleen* 'only' is supported by the fact illustrated in (19) that they can all be pied piped under topicalization of the phrase in their scope. The fact that the finite verb can be preceded by at most one single constituent in Dutch shows that the constituent negator is not an independent part-of-speech (i.e. an adverbial) but part of the topicalized direct object.

- (19) a. Niet het *BOEK* heeft Jan gekocht, maar de *PLAAT*.
 not the book has Jan bought but the record
 'Jan hasn't bought the book, but the record.'
 b. Ook/Alleen de *PLAAT* heeft Jan gekocht.
 also/only the record has Jan bought
 'Jan has also/only bought the record.'

The rating and production experiments in Schoenmakers & De Swart (2019) and Schoenmakers (2020) confirm the special status of examples like (16a). Given that such examples are only interpretable in an appropriate discourse setting, which is not provided in the experiments, it is not surprising that they were rated low and were only rarely produced. The conclusion that scrambling of definite object NPs across the adverb *niet* is obligatory is undoubtedly correct, but the comparison of (16a) and (16b) brings nothing new as they involve completely different syntactic functions of *niet* and completely different syntactic structures, which are sketched in (20).

- (20) a. Jan heeft [[niet het boek] gelezen]. [constituent negation]
 Jan has not the book read
 'It was not the book that he read.'
 b. Jan heeft [het boek]_i [niet [_i gelezen]]. [sentence negation]
 Jan has the book not read
 'Jan hasn't read the book.' (P_A)_S

Although the experiments do no more than confirm Kraak's (1966) original claim that definite objects must precede sentence negation, the rationale for this is still not entirely clear. Schoenmakers & De Swart (2019) resort to a Gricean account for the fact that the examples with constituent negation are rated as less acceptable and are only occasionally produced: "The constituent negation interpretation is pragmatically incomplete, so to speak, if there is no second entity to contrast with the object [...]. This incompleteness is in violation [of] the Gricean maxim of quantity [...], stating that an utterance should be as informative as is required (for the current purposes of the exchange)" (p.137). This tack might very well be on the right track, as the corrective counter-presuppositional focus expressed by constituent negation makes sense only in an appropriate context. Although Schoenmakers & De Swart do not address the question as to why definite object NPs cannot follow sentence negation, it seems reasonable that a pragmatic account is also appropriate in this case: on the assumption that only canonically scrambled definite object NPs refer to activated discourse entities, this might follow from the fact that saying that something did not happen to some uniquely identifiable entity not currently under discussion is simply irrelevant in a developing discourse.

5 Conclusion

This article discusses a number of studies purporting to evaluate the validity of the claims made by a number of theoretically oriented studies based on introspection data. It has been shown that the corpus and experimental studies contain several flaws which make them unfit for this task. This does not imply, however, that corpus and/or experimental studies are inherently incapable of handling the task, as the main problem that I signaled involves the use of temporal adverbs as representatives of the "high" adverbs. Research of the type will provide more reliable results if it satisfies the following requirements: (i) the set of "high" adverbs is to be

restricted to the class of elements previously identified as comment modifiers in Verhagen (1986) in order to ensure that a proper distinction can be made between short and canonical A-scrambling; (ii) the target sentences are to be investigated in sufficient context in order to enable the evaluation of discourse-related claims made in earlier studies like Verhagen (1986), Schaeffer (2000) and Broekhuis (2000/2008).

References

- Aguilar Guevara, Ana. 2014. *Weak definites. Semantics, lexicon and pragmatics*, University of Utrecht: PhD thesis.
- Barbiers, Sjef. 2018. Adverbs in strange places: on the syntax of adverbs in Dutch. *Nederlandse Taalkunde* 23: 57-87.
- Broekhuis, Hans. 2000. Against feature strength. The case of Scandinavian object shift. *Natural Language & Linguistic Theory* 18: 673-721.
- Broekhuis, Hans. 2007. Subject shift and object shift. *Journal of Comparative Germanic Linguistics* 10: 109-141.
- Broekhuis, Hans. 2008. *Derivations and evaluations: object shift in the Germanic languages*. Berlin/New York: Mouton de Gruyter.
- Broekhuis, Hans. 2016. Syntax of Dutch: the data set. *Nederlandse Taalkunde* 21: 297-325.
- Broekhuis, Hans. 2020. Object shift and object scrambling. In *The Cambridge handbook of Germanic linguistics*, eds. Michael T. Putnam and B. Richard Page, 413-432. Cambridge (UK)/New York: Cambridge University Press.
- Broekhuis, Hans & Norbert Corver. 2016. *Syntax of Dutch. Verbs and verb phrases, volume 3*. Amsterdam: Amsterdam University Press.
- Broekhuis, Hans & Marcel Den Dikken. 2012. *Syntax of Dutch. Nouns and noun phrases, volume 2*. Amsterdam: Amsterdam University Press.
- Chomsky, Noam. 1995. Bare phrase structure. In *Government and binding theory and the minimalist program*, ed. Gert Webelhuth, 383-439. Oxford/Cambridge, MA: Blackwell.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Corver, Norbert & Henk van Riemsdijk. 1994. *Studies on scrambling: movement and non-movement approaches to free word-order phenomena*. Berlin/New York: Mouton de Gruyter.
- De Hoop, Helen. 2000. *Otje* en scrambling in het Nederlands. *Tabu* 30: 97-112.
- De Hoop, Helen. 2003. Scrambling in Dutch: optionality and optimality. In *Word order and scrambling*, ed. Simin Karimi, 201-216. Malden, Ma/Oxford (UK): Blackwell Publishing.
- De Hoop, Helen. 2016. Woordvolgordevariatie: theorie versus empirie? *Nederlandse Taalkunde* 21: 265-284.
- De Swart, Peter & Geertje Van Bergen. 2011. *Definiteness and adverb-object order in Dutch*. Ms. University of Groningen/Université catholique de Louvain.
- De Swart, Peter & Geertje Van Bergen. 2014. Unscrambling the lexical nature of weak definiteness. In *Weak referentiality*, eds. Ana Aguilar-Guevara, Bert Le Bruyn and Joost Zwarts, 287-309. Amsterdam/Philadelphia: John Benjamins.
- Den Dikken, Marcel. 1995. *Particles. On the syntax of verb-particle, triadic, and causative constructions*. New York/Oxford: Oxford University Press.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Diesing, Molly. 1997. Yiddish VP order and the typology of object movement in Germanic. *Natural Language & Linguistic Theory* 17: 369-427.

- Diesing, Molly & Eloise Jelinek. 1995. Distributing arguments. *Natural Language Semantics* 3: 123-176.
- Dik, Simon C. 1997. *The theory of functional grammar. Part 1: the structure of the clause*, 2nd revised edition: Mouton de Gruyter.
- Engels, Eva & Sten Vikner. 2014. *Scandinavian object shift and optimality theory*. Basingstoke: Palgrave Macmillan.
- Haegeman, Liliane. 1993a. The morphology and distribution of object clitics in West Flemish. *Studia Linguistica* 47: 57-94.
- Haegeman, Liliane. 1995. *The syntax of negation*. Cambridge: Cambridge University Press.
- Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij & Maarten C. van den Toorn. 1997. *Algemene Nederlandse spraakkunst*, 2nd, revised edition. Groningen: Nijhoff.
- Haider, Hubert, Susan Olsen & Sten Vikner. 1995. Introduction. In *Studies in comparative Germanic syntax*, eds. Hubert Haider, Susan Olsen and Sten Vikner. Dordrecht/Boston/London.
- Holmberg, Anders. 1986. *Word order and syntactic features in the Scandinavian languages and English*, University of Stockholm: PhD thesis.
- Holmberg, Anders. 1999. Remarks on Holmberg's generalization. *Studia Linguistica* 53: 1-39.
- Karimi, Simin. 2003. *Word order and scrambling*. Malden, Ma/Oxford (UK: Blackwell Publishing.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Kraak, Albert. 1966. *Negatieve zinnen: een methodische en grammatische analyse*. Hilversum: W. de Haan.
- Longobardi, Giuseppe. 1994. Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25: 609-665.
- Neeleman, Ad. 1994a. Scrambling as a D-structure phenomenon. In *Studies on scrambling. Movement and non-movement approaches to free word-order phenomena*, eds. Norbert Corver and Henk van Riemsdijk, 387-429. Berlin/New York: Mouton de Gruyter.
- Neeleman, Ad. 1994b. *Complex predicates*, University of Utrecht: PhD thesis.
- Neeleman, Ad & Tanya Reinhart. 1998. Scrambling and the PF Interface. In *The projection of arguments*, eds. Miriam Butt and Wilhelm Geuder, 309-353. Stanford: CSLI Publications.
- Neeleman, Ad & Hans Van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *The Journal of Comparative Germanic Linguistics* 11: 137-189.
- Neeleman, Ad & Reiko Vermeulen. 2012. *The syntax of topic, focus, and contrast. An interface-based approach*. Berlin/Boston: De Gruyter Mouton.
- Newmeyer, Frederick J. 1983. *Grammatical theory. Its limits and possibilities*. Chicago/London: University of Chicago Press.
- Ross, John. 1967. *Constraints on variables in syntax*. Bloomington: Indiana University Linguistics Club [Reprinted as *Infinite syntax!*, Ablex, Norwood New Jersey, 1986].
- Schaeffer, Jeannette C. 2000. *The acquisition of direct object scrambling and clitic placement. Syntax and pragmatics*. Amsterdam/Philadelphia: John Benjamins.
- Schoenmakers, Gert-Jan. 2020. Freedom in the Dutch middle-field: deriving discourse structure at the syntax-pragmatics interface. *Glossa* 5(1):114.
- Schoenmakers, Gert-Jan & Peter De Swart. 2019. Adverbial hurdles in Dutch scrambling. In *Proceedings of Linguistic evidence 2018: experimental data drives linguistic theory*, eds. R. Gattnar, R. Hörnig, M. Störzer and S. Featherston. Tübingen: University of Tübingen.
- Unsworth, Sharon. 2005. *Child L2, adult L2, child L1: differences and similarities. A study on the acquisition of direct object scrambling in Dutch* University of Utrecht: PhD thesis.

- Van Bergen, Geertje & Peter de Swart. 2009. Definiteness and scrambling in Dutch: where theory meets practise. In *Proceedings of NELS 38*, eds. Anisa Schardl and Hannah Morales, 113-124. Amherst: GLSA.
- Van Bergen, Geertje & Peter de Swart. 2010. Scrambling in spoken Dutch: definiteness versus weight as determinants of word order variation. *Corpus Linguistics and Linguistic Theory* 6: 267-295.
- Van der Does, Jaap & Helen de Hoop. 1998. Type-shifting and scrambled definites. *Journal of Semantics* 15: 393-416.
- Vanden Wyngaerd, Guido. 1988. Raising-to-object in English and Dutch. *Dutch Working Papers in English Language and Linguistics* 14.
- Vanden Wyngaerd, Guido. 1989. Object shift as an A-movement rule. In *MIT Working Papers in Linguistics*, 256-271. Cambridge, MA: MIT.
- Verhagen, Arie. 1986. *Linguistic theory and the function of word order in Dutch. A study on interpretive aspects of the order of adverbials and noun phrases*. Dordrecht: Foris Publications.
- Zwart, Jan-Wouter. 1993. *Dutch syntax. A minimalist approach*, University of Groningen: PhD thesis.
- Zwart, Jan-Wouter. 1997. *Morphosyntax of verb movement. A minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer Academic Publishers.