Preposition stranding in development*

Peter Coopmans and Rianne Schippers Utrecht University / UiL-OTS

Introduction

In this paper we discuss experimental evidence concerning an instance of preposition stranding in the language of Dutch 4 and 5-year olds that raises the question whether their grammars are qualitatively different from those employed by Dutch adults. These children's comprehension performance seems to point to a grammar whose stranding mechanism violates standard locality conditions. It raises analytical problems from a grammatical perspective, and conceptual problems from a developmental perspective.

In Section 2 we present the empirical trigger of this research, and the theoretical instruments needed to appreciate the problem. In Section 3 we discuss an experiment and its results on Dutch children's comprehension of preposition stranding, indicating that we are dealing here with a real and truly developmental phenomenon. In Section 4 we place this finding in a broader discussion of apparent violations of locality constraints that might explain these children's preposition stranding behaviour. In Section 5 we propose a more successful alternative in terms of a strict parsing account based on the head-driven model of Pritchett (1992). We conclude that on such an account this phenomenon will evoke less curiosity, since it can be shown to be ultimately reducible to an independently attested feature of preposition stranding development in Dutch.

The facts

The trigger of this research project is the observation the first author made when confronted with the following question by his 3-year old daughter.

(1) Wie heeft van Cas dit gekregen? who has from Cas this received

(Fenna 3;5)

The order of words forming this question should not trigger any surprise were it not for the fact that its intended meaning was exactly the opposite of how adults interpret this sentence.¹ Dutch grammar only allows the representation in (2a), with the subject role assigned to the question word *wie*. We will call this the *Wh-sub (subject)* reading. But this is not what the speaker had in mind. She asked this question with the interpretation represented in (2b), where *wie* is the source, and *Cas* is the recipient. We will call this the *Wh-pob (prep. object)* reading.

- (2) a. Wie heeft [__ [van Cas] dit gekregen]]'Who received this from Cas?'b. Wie heeft [[van __] Cas dit gekregen]]'Who did Cas receive this from?'
- (2b) seems to be a clear case of preposition stranding. This process is available in Dutch, albeit in a limited form. As has been extensively documented by van Riemsdijk (1978), preposition stranding in Dutch is possible only with so-called '[r]-pronouns', as shown in (3). (3a) shows movement of the entire PP to the front of the interrogative clause, like its pied-piped English equivalent *Of what did you think?* (3b) shows that the interrogative pronoun *wat* cannot strand the preposition *aan*. To derive the legitimate Dutch counterpart of *What did you think of?*, the wh-element needs to be transformed into a [r]-pronoun, in this case *waar* (3c). The version in (3d) is the alternative of (3a), again a case of pied piping, but with the [r]-variant preceding the preposition.

This combination of forms, among other things, led van Riemsdijk to propose that only [r]-pronouns can strand the preposition, and that movement out of the PP is made possible by an internal escape route via the specifier position. The [r]-form is an overt reflex of this (temporary) landing in spec-PP. A central part of this account was that PPs are bounding nodes for extraction for everything c-commanded by the P. This legitimises spec-PP as an escape hatch, in line with earlier work on locality constraints in the 1970s. Here we will simply be assuming van Riemsdijk's general account of Dutch grammar as having a limited form of preposition stranding through spec-PP, overtly reflected by a morphologically identifiable class of [r]-pronouns.

Dutch children take a while to master the fact that preposition stranding is limited to this specific class of pronouns. Initially, they seem to have a more liberal

English-type grammar, with no restriction to particular pronouns whatsoever. Some examples are given in (4), where none of the italicised forms are [r].

(4)	a.	Weet je welke deze is van?	(Sara 4;1)
		'Do you know which this is from?'	
	b.	Oh, wat is dat nou weer voor?	(Tim 3;5)
		Wie is dat voor?	(Tim 3;4)
		'What/who is that for?'	
	c.	Weet je nog meer wat ik over heb gedroomd?	(Sarah 3;11)
		'Do you know what I have dreamt of?'	
	d.	<i>Da(t)</i> ben ik naar heen geweest.	(Daan 3; 3)
		'That have I been to'	
	e.	Dat gaf je guldens aan.	(Tinke 3;9)
		'That you gave guilders to.'	

The ability of Dutch children to produce such structures as (4) does not necessarily require us to postulate a grammar that is different from the adult target. If we assume that the stranding they perform is made possible by an intermediate step via spec-PP, the only difference is a fairly superficial one: children will need to learn that in Dutch the feature that triggers such PP-internal movement goes hand in hand with an obligatorily overt reflex [r].²

If we take this to be a general feature of a Dutch child's developing grammar of preposition stranding, then what makes the observation in (1) so special? Isn't this characteristic of the general case? The answer should be sought in the position of the phrase from which stranding seems to have taken place: a PP to the left of the subject.³ Its position suggests that the PP is located outside the immediate vicinity of the lexical verb, turning it into an island for subextraction if overt movement needs to obey a locality constraint of the CED type:

(5) Condition on Extraction Domains (Huang 1982) A phrase β cannot be extracted from a domain α if α is not properly governed.

A more specific analysis of (2b) is given in (6). Here the PP adjoined to TP would surely qualify as an α according to the definition in (5), and hence as an island.

(6) Wie heeft $[_{TP} [_{PP} \text{ wie } [\text{van wie}]] [_{TP} \text{ Cas dit gekregen}]]$

In her production of a question like (1), on the intended interpretation, the child seems to be violating a universal condition on overt movement, posing a more serious problem from an acquisition perspective than the lack of language-specific knowledge of the [r]-requirement. Is this phenomenon observed more generally? If it is, this developmental stage requires a deeper explanation. In the next section we will report on how we tested the generality of this phenomenon.

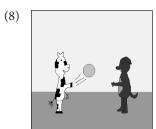
3. The experiment

In order to find out whether Dutch children accept a form of illegitimate stranding as illustrated in (2b), we devised a comprehension test for 4 and 5-year olds, consisting of two tasks: one part following a standard truth-value judgement (TVJ) design, the other following a questions-after-stories (QaS) design (cf. McDaniel et al. 1996). The tasks were administered in two separate sessions. The general idea was to confront children with a situation in which in principle both interpretations (*wh-sub* and *wh-pob*) could be associated with a question like (1). This would ensure that if they had a grammar that would allow a form of stranding like in (2b), they could in principle have two possible answers to such a question. A QaS design might help in determining a possible preference, while a TVJ design could determine whether they would categorically allow or reject the stranding displayed in (2b). Because a TVJ design is more restrictive than a QaS design, the TVJ task was carried out in the first session and the QaS task in the second. The children were presented with four stories in total, which were equally divided over the two sessions. Each story consisted of two test sentences, a filler item and a control item. And each session was started with two warm-up items, introducing the main protagonists of the stories: cow, dog and chicken.⁴ An example of a story is given in (7), with the Dutch test sentence in italics.

Three actions in this story are relevant to the target question. Two of them correspond to the two interpretations under discussion: cow throwing the ball to dog corresponds to the *wh-sub* interpretation and dog throwing the ball to chicken corresponds to the *wh-pob* interpretation. The third action is there for counterbalancing and to make each protagonist equally likely as an agent as well as a receiver. We varied the orders of presentation, to prevent any unwanted effects on preferred outcome for either the *wh-sub* or the *wh-pob* interpretation: each story thus had two test sentences, each with a different order, and each story had two versions, where the orders for the test sentences were reversed. All stories were supported by accompanying pictures, the one in (7) by the pictures in (8). Each picture was introduced individually, so that the children would first only see the leftmost picture, then the leftmost and the middle pictures, and finally they would see all three pictures. All pictures then remained available during the trial.

(7) They all go outside to throw the ball to each other. They form a circle and throw the ball to each other. Cow throws the ball to dog, dog throws the ball to chicken and chicken throws the ball to cow.

Wie heeft naar hond de bal gegooid? who has to dog the ball thrown?



(0)





Because we were testing to see whether children would allow for an illegitimate form of preposition stranding, we had to make sure they allowed for legitimate preposition stranding in the first place. The control items were designed to do exactly that. The general idea was to present children with a question that could be interpreted with and without the stranded preposition, leading to two different interpretations. Children who did not allow for preposition stranding would not know what to do with a stranded preposition and hence ignore it. On the basis of the answers to the control items we could then see whether the child interpreted the stranded preposition, and thus determine whether the child allowed for legitimate preposition stranding.

58 children were tested: 29 4-year olds (range 4;0–4;11; mean age 4;5), and 29 5-year olds (range: 5;0–6;2; mean age 5;4), as well as a control group of 9 adults, all non-linguists.⁵ The children performed adultlike on both the fillers and the control items. The results of their behaviour (percentages correct/wh-sub reading) on the individual test sentences are given in (9). In this table T stands for TVJ and Q for QaS. The numbers after these labels refer to the test sentences. So T-1 refers to the first test sentence in the TVJ design, T-2 to the second, etc. The last two columns contain the overall results for the TVJ and the QaS components of the test.

(9)	(9)										
	T-1	T-2	T-3	T-4	Q-1	Q-2	Q-3	Q-4	TVJ	QaS	
4 y.	50	50	64	36	34	69	66	52	50	55	
5 y.	61	32	68	50	55	45	69	55	53	56	
adults	100	100	100	100	100	100	100	100	100	100	

Percentages 'correct', i.e. wh-sub, judgments on TVJ-target questions/answers, and wh-sub answers to QaS target questions. No significant difference between 4 and 5-year olds. On TVJ: t=-.606; df=53; p=.54; On QaS: t=-.151; df=55; p=0.88, equal variances assumed. Significant difference between groups of children and adults. On TVJ: Kruskal Wallis $\chi^2=19,457$; df=2; p<0.001; On QaS: Kruskal Wallis $\chi^2=8,315$; df=2; p<0.001.

First, the results in (9) clearly show that the 4 and 5-year olds performed significantly differently from the adults. The children allow illegitimate stranding in about 50% of the cases, suggesting that the test sentences were indeed ambiguous to them. Secondly, there was no statistically significant difference between the

two groups of children. They seemed to equally accept this form of illegitimate stranding. That most children did indeed allow both the *wh-sub* and the *wh-pob* interpretation can be concluded from the fact that no child accepted the *wh-pob* interpretation only, and that there were four children who had a *wh-sub* interpretation only. These four children were equally distributed over the 4 and 5-year olds, confirming the general pattern we see in (9).

4. Grammatical development?

The experimental results confirm our original suspicion that we are dealing with a real developmental phenomenon. Dutch 4 and 5-year olds go through a comprehension stage during which they not only allow a regular non-[r] wh-pronoun to strand a preposition, but also seem to be able to do so in an environment that should not permit stranding in the first place (assuming that the fronted PP to the left of the subject is essentially an island). This modification follows the standard position in the literature on locality. If the fronted PP qualifies as a CED island, then the question arises whether there is any specific reason why it should be less of a bounding domain in children's grammars.

Movement of the PP to a position to the left of the subject induces a freezing effect of the type studied by Wexler and Culicover (1980). However, certain comparable instances suggest that this reasoning may be too strict. Coopmans (1988), for example, discussed instances of moved phrases which are relatively transparent for further extraction, and argued that a distinction should be made between base-generated and derived adjuncts, the latter semi-transparent for subextraction. These include postverbal definite subject arguments in Italian and scrambled indefinite objects in Dutch. From a representational perspective, the derivational history of these arguments should have no effect on the strength of their bounding nature (following the CED), but it does matter. Turning them into CED-eligible environments by adjunction nevertheless gives them less of an island character than if they had been base-generated as such.

Similarly, Lasnik and Saito (1992:111) showed that topicalising an argument in English by adjoining it to IP does not induce a strong island effect for subextraction. The difference is illustrated by the following examples.

```
(10) a. Who do you think that [John bought [pictures of __]]
b. *Who do you think that [[pictures of __] are on sale]
c. *? Who do you think that [[pictures of __] [John bought]]
```

(10a) and (10b) show the standard object-subject asymmetry in bounding domains. Extraction from a complement is fine, not so from a subject. (10c), with

the object argument topicalised, should be as bad as (10b) but appears 'reasonably acceptable'.

We could follow this line of reasoning in our treatment of (2b)/(6), and argue that extraction from such fronted PPs is indeed reasonably acceptable. This will be problematic, however, in view of the observation that a stranded preposition as the sole element in topic position in English is always ill-formed, as is shown by (11). Whatever accounts for the ill-formedness of (11) would in principle also hold for (2b).

(11) *Who do you think that [[to __][John gave a book]]

In addition to the kind of empirical problem illustrated here by English topicalisation, there is the general acquisition problem of why Dutch children would behave differently from adults. Either fronted PPs are transparent, and the children's behaviour, but not the adults', would be in accord with linguistic theory, or the opposite holds: fronted PPs are islands, adults would obey them, and children would not. Why would children start out with a softer locality constraint, and then slowly learn to obey a stronger one? The relevant positive evidence would be lacking from the input. In fact, neither scenario permits a reasonable explanation for why a developmental change in grammar could be triggered as a result of which children would no longer accept (2b) as a possible structure for the question in (1). This general acquisition problem suggests to us that we should look for an explanation outside the grammar proper.

5. A parsing account

The discussion so far has centred on two empirical observations: (a) Dutch children are more liberal than adults in allowing non-[r]-pronouns to strand prepositions; (b) Dutch children are more liberal in allowing stranding from topicalised (argument) PPs. The account that we will propose here assumes that fact (a) is in part responsible for the phenomenon described in (b). We will put forward the view that these are not separate phenomena, and that the occurrence of (b) is just a side effect of (a). Dutch children simply have to learn that the pronominal elements that can strand a preposition are restricted to the class of [r]-pronouns. Once they have learned this, we think that (b) should no longer be attested either. We will sketch a parsing account here based on this line of reasoning. The account assumes a strict interpretation of the head-driven parsing model of Pritchett (1992).

In Pritchett's model syntactic processing is driven by local application of grammatical principles. Specifically, the structure-building process is driven by theta attachment: "The theta criterion attempts to be satisfied at every point

during processing given the maximal theta grid." (Pritchett 1992:12). For the general parse at clausal level, it is the lexical verb which drives the structural attachment of arguments. This means that arguments will have to be stored until such a theta-assigning verb is encountered. In a verb-final language like Dutch a number of arguments will remain unattached until the end of the clause delivers the necessary source of thematic information (for discussion, see Mulders 2002).

Let us analyse the child's parse of (1) step by step as in (12), assuming that she lacks the [r]-constraint. Steps (i) and (ii) should be straightforward. *Wie* is analysed as a Q-nominal, followed by a verb. Note that under a strict interpretation of head-driven parsing, encountering *hebben* does not lead to any further attachment on the assumption that such auxiliary elements do not assign thematic roles by themselves.⁸ Everything so far is kept in store. The third element is the preposition *van* (iii), which requires a complement. In other words, theta attachment is invoked.

(12) Wie heeft van Cas dit gekregen?

i.	wie	$N_{[+Q]}$
ii.	wie heeft	$NP_{[+Q]}V$
iii.	wie heeft van	$NP_{[+Q]} V P$

There are now two possibilities. Let us follow one until the end, and then return to the second possibility. A complement to *van* immediately presents itself when *Cas* is parsed, and a PP can be built. This is step (iv). The next step involves the parse of *dit*, another nominal element (v). None of the arguments will be attached until the participle *gekregen* is parsed, providing the necessary theta roles. This is how the child can build the structure for the *wh-sub* reading.

iv. wie heeft van Cas
$$\begin{array}{ll} NP_{[+Q]} \ V \ [P \ N] \\ v. & \text{wie heeft van Cas dit} \dots & NP_{[+Q]} \ V \ [_{PP} \ P \ NP] \ N \dots \end{array}$$

The alternative possibility which the child has at her disposal is to continue from step (iii) by postulating an empty complement to satisfy theta attachment. The interpretation of this empty complement is licensed by the Q-element encountered earlier, forming the basis for an operator-variable structure. We will represent this complement by t, traditionally for trace. The child can build such a structure because her grammar allows the link between a non-[r] wh-pronoun and a gap immediately following a preposition. This is shown in step (iv'). Step (v') then analyses Cas as a nominal element, and (vi') will do the same to dit. Like before, no attachment of any of these arguments (with the exception of the wh-trace in complement position) will take place until the parse of the participle gekregen. This is how the child can build the alternative representation for the wh-pob reading.

This alternative parse is not entertained by adults, because in their system there is no grammatical basis for the link between *wie* and *van* ..., due to the [r]-constraint. They will go directly for the parse of the *wh-sub* reading (steps iv–v) and no other.

As long as Dutch children lack a firmly engrafted [r]-constraint, they will show comprehension performance that reflects ambiguity of structure assigned to questions like (2). It is important to emphasize here that by devising the experimental conditions in the way we did, we tricked the children into displaying the effect of illegitimate stranding, seemingly creating an island violation. Yet, basic island-hood does not play a role. The difference between children and adults only resides in that the grammar of children still allows for a relation between wie ... [van ...], an option that they can then automatically use. The development, then, is not a grammatical one, involving, say, a softer to a stronger CED constraint.

If this approach is on the right track, a few further remarks are in order. First of all, we have capitalised on the fact that no attachment is possible until the lexical verb is analysed, and we have crucially assumed that parsing of the auxiliary *hebben* does not trigger theta attachment. This predicts that if we had used a finite form of the lexical verb *krijgen* 'receive' in V-2nd position, immediately following *wie*, we would have enforced attachment of the wh-element before the analysis of the preposition, and the results might have been much more adultlike: *Wie kreeg van Cas dit (boek)*? 'Who received from Cas this (book)?'. This is an unforeseen effect of the experimental input. Still, it provides a perfectly testable hypothesis to weigh the plausibility of our parsing approach. We will take this up in a further experiment.

Related to this point is the criticism that a parsing effect should not show up so robustly in an offline experiment. After all, the children probably heard the lexical verb before answering the question, or passing their judgement. Still, they would have started their analysis, and in a number of cases we noted that they responded to the QaS questions with a full PP answer, as if they had initially parsed: [Van wie] heeft Cas ... 10 This is something that we will need investigate more closely in a follow-up experiment as well.

An adult counterpart of (2b) would be the ungrammatical question in (13). This raises the question how this can be accounted for under the parsing approach described here. Let us look at it in more detail.¹¹

Steps (i–iii) are no different from those in (12). The crucial question is whether in the continuation the complement of the preposition is necessarily occupied by the next element encountered, *Cas* (step iv above), or whether the option exists of meeting theta attachment by postulating an empty complement, to be linked with the Q-element *waar* (step iv' above). The difference between (13) and (2b) is that *waar* as an r-pronoun is ambiguous between an argument (what) and an adjunct (where). At step (iv) of the parse, this ambiguity cannot be resolved, and further analysis or attachment driven by *waar* needs to be suspended. Step (iv) thus needs to be continued with attachment of *Cas* as the complement of *van*, and a PP can be built. This is the actual step (iv). Step (v) then proceeds as before.

```
iv. ?waar heeft van NP_{[+Q]} V [P t] ambiguity of waar: \rightarrow suspend analysis \sqrt{\text{waar}} heeft van Cas NP_{[+Q]} V [P N] v. waar heeft van Cas dit ... NP_{[+Q]} V [_{PP} P NP] N
```

When the lexical verb is analysed, and the theta roles are distributed (so that the various arguments can be attached), it turns out that the postulated analysis is one argument short. *Cas* should be projected onto the subject function, but has instead been incorrectly parsed as the complement of the preposition, resulting in a theta violation. This account thus sheds a theta-theoretic perspective on a subcase of a traditional bounding phenomenon.

6. A problem from English

We will conclude with a final remark about English, where, if we take (14a) as the interrogative counterpart of the topic construction in (14b), a similar construction can be set up. While it would be interesting to see how children would respond, the adult judgements already call for an explanation.

```
(14) a. *Who will [from __] John receive a present tonight?b. [From Mary] John will receive a present tonight.
```

Perhaps (14a) is ill-formed for additional reasons, in particular the very fact that topics themselves generally create islands for elements that want to move across them, and movement of the auxiliary *will* in (14a) would be constrained by the presence of topicalised PP. This may have to do with the highly restricted nature of topicalisation in English. We have seen above that the account for English requires an independent explanation for the prohibition against stranded prepositions in topic position, and in this respect (14a) would be no different from (11). Why this should be so remains unexplained.

Notes

- * We thank Iris Mulders and two anonymous reviewers for very useful criticism, and Jacqueline van Kampen for providing some production data (from the van Kampen and Schlichting corpora).
- Actually, she used *heb* rather than *heeft*, but for ease of analysis and presentation we are ignoring this typical feature of Dutch child language, since it is irrelevant for the point we will make here.
- 2. Stated as such, we are treating what is essentially a subset problem (the child has to acquire a more restrictive system) as a problem of lexical learning. There is clearly much more to say about this, perhaps also from a wider perspective (cf. van Kampen 1997), but we cannot do so here for reasons of space.
- **3.** The other order can be found in Dutch child language too: *Wie heeft Cas van dit gekregen? Such examples essentially reduce to those in (4), i.e. liberal stranding from a complement PP.
- 4. Space limitations prevent us from giving full details of the test here, the various conditions and further quantitative results, but see Schippers (2007).
- We also tested nine 3-year olds but excluded them from further analysis, since they failed the control conditions.
- **6.** For recent discussion incorporating derivational history into the character of even subject islands, see Chomsky (2008).
- 7. This is the hypothesis that we can formulate on the basis of our finding. Clearly, we have not tested for any correlation, since this was not the goal of our experiment. It was set up simply to see whether the anecdotal production evidence was somehow mirrored in comprehension performance.
- **8.** This will extend to all variants of *hebben* 'have', including the possessive, if these can also be reduced to an auxiliary structure.
- 9. At this stage of our investigation, we have no way of telling whether one of both options could perhaps be considered default option.
- 10. To add some anecdotal evidence, the first author has independently asked some 7–11 old children to literally repeat what he asked; some of them incorrectly responded with "Van wie heeft ...'
- 11. A reviewer points out that the example we are using here for illustration may not be entirely appropriate because the alternative *Waar heeft Cas dit van gekregen?* may not be felicitous when he obtained a book, and sounds more natural if he contracted a disease. A better example, but further removed from our original example, would be *Waar heeft voor [___] Cas dit gekocht? ('Where has ... bought').
- **12.** Alternatively, under theta attachment, analysing *waar* as an adjunct would result in fewer local violations of the theta criterion, and would be the preferred analysis (Iris Mulders pc).

References

- Chomsky, N. 2008. 'On Phases'. Foundational Issues in Linguistic Theory. ed. by R. Freidin, C. Otero and M.L. Zubizarreta, 133–166. Cambridge, Mass.: MIT Press.
- Coopmans, P. 1988. 'On Extraction from Adjuncts in VP'. *Proceedings of the Seventh West Coast Conference on Formal Linguistics*. ed. by H. Borer, 53–65. Stanford Linguistics Association.
- Huang, C.T. J. 1982. Logical Relations in Chinese and the Theory of Grammar. PhD diss., MIT.
- Kampen, J. van. 1997. First Steps in Wh-Movement. PhD diss., Utrecht University.
- Lasnik, H. and M. Saito. 1992. *Move α. Conditions on Its Application and Output.* Cambridge, Mass.: MIT Press.
- McDaniel, D., C. McKee and H. Smith Cairns, eds. 1996. *Methods for Assessing Children's Syntax*. Cambridge, Mass.: MIT Press.
- Mulders, I. 2002. Transparent Parsing. Head-Driven Processing of Verb-final Structures. PhD diss., Utrecht University.
- Pritchett, B. 1992. Grammatical Competence and Parsing Performance. Chicago: Chicago Univ. Press.
- Riemsdijk, H. van. 1978. A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases. Lisse: Peter de Ridder Press.
- Schippers, R. 2007. A Certain Case of Preposition Stranding in Dutch First Language Acquisition. Internship report, UiL-OTS.
- Wexler, K. and P. Culicover. 1980. Formal Principles of Language Acquisition. Cambridge, Mass.: MIT Press.