An Acquisition Procedure for Category Assignment
The Case of Dutch P⁰

Jacqueline van Kampen & Rianne Schippers
UiL OTS, Utrecht University

1. The category P: Prepositions and particles

The present note returns to a set of distributions for the Dutch category P, once referred to by Van Riemsdijk (1973, 1978; see also Ruys, 2008) as the Dutch "P-soup". The point of attention here will be how the relevant P-subcategories are identified and acquired in child language.

A core set of preposition-like lexical items in Dutch (like *aan* ('on'), *in* ('in'), *op* ('up'), *uit* ('out')) appears quite early in two-word utterances of child language as minimal predicates. See some attested examples in (1).

(1) a. moet aan   b. schoen aan   c. wil   op   d. pop op   e. dit uit
        must on   shoe   on   wanna   up   doll up   this out

When, in a next stage, the syntactic context expands in nouns and verbs, the prospective items P appear in two rather different environments [P + DP] (2a) and [P + V] (2b). Emonds (1985) distinguished the two types of P as transitive P versus intransitive P. The same elements P can be used both ways but with different semantic effects and with further differences in syntactic distribution as well, as will be elaborated in section 3. Both P's, the one associated with a DP (2a) and the one directly contiguous with the Verb (2b), are selected by the predicate head V (*halen* ('take') and *spugen* ('spit').

(2) a. wil   [uit (de) doos]   halen
        wanna   out (the) box   take
        'I wanna take (this) out (the) box'
    b. ga   [uit spugen]
        go   out spit
        'I am going to spit (this) out'

We will distinguish the two subcategories of P as Prep (preposition) and Prt (verbal particle). The fact that the same items may appear either as Prep or as Prt given the appropriate verb requires an historical explanation (Blom, 2005). That explanation will plausibly be rooted in the acquisition procedure for P. Our present concern is a procedure in language acquisition that identifies the two P subcategories and subsequently stores them as such in the lexicon. It may seem to be a trivial matter of repetitive contexts in the input, but a tricky complication arises. See below, sections 4-5. Section 6 will bring us back to the influence of child language acquisition on language change.

Section 2 will offer a look at the (sub-)procedure for lexical category assignment in general. The following should be kept in mind. Any acquisition procedure must assign syntactic categories to lexical items (and phrases). The lexicon is language-specific. Hence, the acquisition procedure for category assignment has to apply to language-specific items, item by item, given some decisive language-specific evidence for category determination. It does not really matter whether one assumes in addition that all syntactic categories are wired into the human brain a priori, as in Pinker (1984). Discriminating evidence from actual input strings is needed anyway.

2. Systematic input reduction and category acquisition

Categorial as well as Minimalist style of grammars are based on the assumption that all combinations of elements or phrases are binary, that all phrases have endocentric heads, and that this allows a bare phrase analysis. That is, no distinction needs to be made between category labels for phrases and category labels of their terminal heads. The bare phrase assumption bears a promise for acquisition theory. Once the appropriate labels for the terminal elements are established and stored lexically, learning the grammatical system may run like a well-guided parsing automatism. By consequence, the bare phrase approach may direct the attention to the procedure that determines the appropriate category assignment for lexical items. We assume that the acquisition procedure needs to assign syntactic category labels to lexical items by means of a language-specific context. Our language-specific approach to category assignment does not ignore that the adult system may get disturbingly complex, but we assume that there is a way in. As has been observed by Jakobson (1942) and applied for syntax by Lebeaux (2000), there is an effective code-breaking procedure for language acquisition. The language acquisition procedure must start with a massive and systematic input reduction. Fortunately, input reduction is the hallmark of early child language. Full-fledged adult sentences are reduced to binary combinations of two content words that were acquired in a previous pre-syntax stage as single items. At the very start, all (or almost all) grammatical markings are left out as not yet interpretable. We take it that the reduction, - leave out any element that cannot yet be interpreted -, is not based on a priori information, but rather on the ignorance of the acquisition procedure about the system that is to be decoded. The first distinctions are made and acquired within simplified systems. See Van Kampen (1997, chapter 2; 2009) for references and a more elaborated exposition. One of the binary constructs in early child language is a situational deictic element followed by some stressed pre-categorial content word as a commenting element. See (3).

(3) a. die hier   (that here)   (ADV)
    b. die op   (that on)   (P)
    c. die mooi   (that nice)   (A)
    d. die pappa   (that daddy)   (N)
   e. die slapen   (that sleep)   (V)

The later syntactic categories of these commenting elements are added between brackets, although the initial system as manifest in (3) does not yet offer any ground for making such category distinctions. This view does not assume that the learner has an a priori syntactic or semantic distinction between categories. Any of the comment parts in (3) can be understood as reflecting a state of affairs or an event. Nevertheless, the relevant syntactic categories will soon be acquired. We assume that lexical items are routinely categorized as V(erb), and as such lexically stored, when they fit a predicate marking by some I^o (copula, modal, aspectual, light verb, inflection). These I^o markings are stereotype highly frequent additions in the input and sensitive to the category of the predicate item. In the same way lexical items get categorized and stored as N(oun) when they fit the context of reference marking by D^o, see the categorization rules in (4).

(4) a. X → <+V> / I^o <+finite> ___
    b. X → <+N> / D^o ___

Morpho-syntactic bootstrapping of lexical categories is outlined in Maratsos and Chalkey (1980) and reappears in various computational approaches to distributional learning (Cartwright & Brent, 1997; Redington, Chater & Finch, 1998; Mintz, 2003, among others).

Somewhat before their third birthday most children have solved the V^o/I^o and the N^o/D^o categorization, as Figure 1 (taken from Van Kampen, 2004) exemplifies. The graphs represent the acquisition of I(nflection)-marking and D(eterminer) marking for the child Sarah (Van Kampen corpus in CHILDES).

Figure 1   Dutch Sarah: Acquisition of I(inflection) marking and D(eterminer) marking
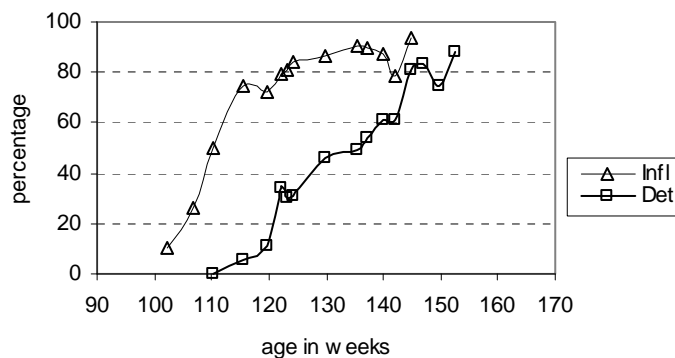
The graph at the left reflects the percentage of sentences where the originally primitive comment gets I^o (inflection) marked and thereby turns its comment phrase into a grammatical predicate IP. Inflection differences (e.g. ± copula) separate the verbal and non-verbal predicates. The systematic reference marking by D^o (determiners), represented by the graph at the right in Figure 1, clearly follows the predicate marking by I^o. Similar acquisition graphs have been made for three other Dutch children (Schippers to appear). We will not consider here how the acquisition procedure may succeed in interpreting the recurrent markings as the respective functions of predication and reference. See for a somewhat more detailed analysis Evers and Van Kampen (2008) and Van Kampen (2009). Leaving that aside, there is at least the simple outcome of the lexical categorization procedure stated in (4) for all nouns and verbs. The rules in (4) are not meant as phrasal rules, but as acquisition rules that assign a syntactic category given a binary phrase. The function of the phrase (predication, reference) must have been recognized as reference pointing and predicate attribution. The I^o and D^o markings themselves are highly language-specific but also highly frequent in the input. Context rules as in (4) should also differentiate prepositions and particles (and adjectives and adverbs).

Since prepositions Prep and verbal particles Prt share the same phonological forms, a procedure like in (4) must identify different category assignments (P_{Prep}, P_{Prt}) for the same phonological form. One may realize though that such complications hold for category assignment in general and that the categorial differentiation of such forms will follow from further syntactic articulation. Local contextual evidence in the string is the key factor. For example, content items like *sleep*, *drink*, *walk* come out as <+V> (context I^o <+finite>) or as <+N> (context D^o). The distinction between nouns and verbs arises when the appropriate markings <+I> or <+D> are systematically added.

As an anonymous reviewer remarks, children may be said to start with predicative particles (3b) and develop prepositions only later (Tomasello, 1987; Morgenstern & Sekali, 2009). The present approach operates by means of context-sensitive acquisition rules and sees this slightly different. The commenting element *op* in (3b) is not yet a particle. There is at that moment in early child language not a distinction between Prt and Prep. The grammatical context ([X + V] (particle) versus [X + DP]) (preposition) has to be established first before the acquisition procedure will differentiate between the two. We will now argue along these lines.

3. The P subcategories in Dutch

Like many other P-elements in Dutch, *aan*, *in*, *op*, *uit* ('on', 'in', 'up', 'out') can be selected as a Prep or as a Prt, even given the same verb. See the examples in (5) for the Prep and (6) for the Prt. The words in bold indicate each time the most likely targets for the neutral predicate stress.

(5) a.   over   het **boek**   schrijven
            about   the book   write
            'write about the book'

b. er-over **schrijven**
   there-about write
   'write about it'
c. er … niet over **schrijven**
   there … not about write
   'not write about it'

(6) a. het boek **over**-schrijven
   the book over-write
   'copy the book by hand'
b. het **over**-schrijven
   it over-write
   'copy it by hand'
c. het … niet **over**-schrijven
   it … not over-write
   'not copy it by hand'

The b- and c-variants in (5) and (6) show that there is a major difference between Prep and Prt concerning object pronouns. The Prep object must be present and when pronominalized it is an oblique weak pronoun *er/d'r* ('there'), or a full oblique pronoun like *daar* ('there'), *hier* ('here'), *waar* ('where')). The [Prt + V] object (if present) is an accusative weak pronoun *'t/het* ('it'), or another accusative pronoun like *dat/die* ('that'), *dit/deze* ('this') or *'m/hem* ('him'), *d'r/haar* ('her'). The dots in the c-variants show how the pronominal elements need not yield strictly local evidence for a Prep versus Prt distinction. The pronoun *er* in (5c) is not adjacent to the Prep and the pronoun *het* in (6c) is not adjacent to the [Prt + V]. The pronouns shift in principle to the left and out of the predicate.

The adjacency of the Prt and the V within the complex [Prt + V] suggests a kind of lexical compound item with internal syntactic structure (Neeleman & Weerman, 1993; Snyder, 2007). It must be noted, though, that Dutch auxiliaries, semi-auxiliaries and causatives may (but need not) intervene between the Prt and its lexical selector V, suggesting that the Prt, although V-adjacent and idiomatic in meaning, is still a separate word. See the two variants in (7).

(7) a. het boek niet [gaan **over**-schrijven]
   the book not go over-write
   'not going to copy the book'
b. het boek niet [**over** gaan schrijven]
   the book not over go write
   'not going to copy the book'

As (7b) shows, the Prt has some distributional leeway within the array of predicate heads. The adjacency between particle Prt and some label V, though, continues to be maintained. The same holds for the predicate stress on the Prt. Stress is maintained irrespective of the Prt's position. Moreover, a verb V may, and even must be, extracted from this predicate construct [Prt + V] if the V is finite and has to serve as the illocutional head in root sentences (Rizzi's "Force" operator). The relevant rule, Den Besten's (1973) V-second rule, moves the finite verb to C^o. When the V-second rule extracts the finite verb from the predicate construct [Prt + V], it strands the particle Prt in a predicate final position, [Prt + t_V]$_V$. See (8).

(8) hij schrijft$_{<+V, +finite, +C>}$ het boek niet [**over** t_V]$_V$
   he writes the book not over
   'he does not copy the book by hand'

Another distributional fact is brought in by the preposition phrase [Prep + DP]. When the DP takes the form of an inanimate pronoun, it gets marked as inherently oblique (*r*-marked). The *r*-marked pronouns move to the left and obligatorily strand the preposition. See (5c) repeated as (9).

(9) [er]$_{DP}$ …niet   [over t$_D$]$_{PrepP}$ **schrijven**
    there … not   [about]   write
    'not write about it'

The shifted oblique pronoun belongs to the Prep construction. A pronominal object of the [Prt + V] construction takes the accusative case. The accusative pronoun shifts to the left as well. See the accusative pronoun *het* in (6b/c).

Taking them together, we see the following three differences in *order*, *stress* and *selection* between Prep-status and Prt-status of a lexical item P.

(10) a. The Prt is predicate-final and V-adjacent.
        The Prep is D-adjacent, unless it is stranded by *r*-pronoun shift. If stranded, Prep will just like the Prt appear V-adjacent or predicate-final.
    b. The Prt gets the predicate stress in a predicate-final array of predicate heads.
        The Prep generally does not get such stress and does not enter the array of predicate heads.
    c. The pronominal object in a Prt construction is an accusative (and shifts).
        The pronominal object in a Prep construction is an oblique (and shifts).

It should be noted that the stranded Prep sometimes does bear the stress, especially when the adjacent verb is a light verb, as in *er in doen* ('in there put'), an issue still to be analyzed

The question now is at which point the acquisition procedure can keep track of all these distinctions. We therefore reconstruct the acquisition procedure for category assignment as an approximating procedure, a stepwise decoding. The fact that mistakes are relatively speaking rare and often short-lived (Snyder, 2007), testifies that the acquisition procedure is highly effective. Nevertheless, mistakes, and even systematic mistakes, do occur for child language in a predictable fashion. It is crucial, though, to focus on the right data in the right period of acquisition. If so, the acquisition of the two P subcategories in Dutch reveals that the child's input reduction (object drop) leads to an acquisition error (stranded Prep analyzed as Prt).

4. The ambiguity of P between Prep and Prt

Once the categories I^o and D^o have been established as regular structure builders, a provisional category decision can be made for the fixed elements P that are clearly neither <+V>/<+I>, nor <+N>/<+D>. There is a subset of P-elements that is regularly followed by a DP. Stress falls in the direction of the DP and away from the P. These elements P may be captured by the category assignment in (11).

(11)   P → Prep / [___ DP]

The provisional category P was present in early syntax among the commenting elements in (3). Restrictions in predicate marking separate these early P's from <+V>, <+N>, and <+A>. Licensing contexts for the single predicative P are the copula, aspectuals, modals, light verbs and a set of more specific lexical verbs, known as the particle verbs. We propose the rule in (12).[1]

(12)   P → Prt [___ V]$_V$

There is a difference between the homophonous elements P in (11) and (12). The phrasal stress in binary structures falls on the discriminating element, i.e. in the direction of the non-head. As verbal

---

[1] This simplifies the discussion by leaving out adjectives and adverbs used as V-adjacent secondary predicates, i.e. *opendoen* ('open do', to open), *wegzetten* ('away put', to put away). These do at that moment in child language not yet differ significantly from particles. They are not considered here, because they have no prepositional counterpart.

particles, stress falls on the Prt and away from the context V. The V in [Prt + V] qualifies as the head, projects, and is not stressed, see (13a). As prepositions, phrasal stress falls in the direction of the DP and away from the Prep. The Prep in [Prep + DP] qualifies as the head, projects, and therefore it does not get the phrasal stress, see (13b).

(13) a.   V   b.   P

                 s   w   w   s

               P   V   P   **DP**

              stress   stress
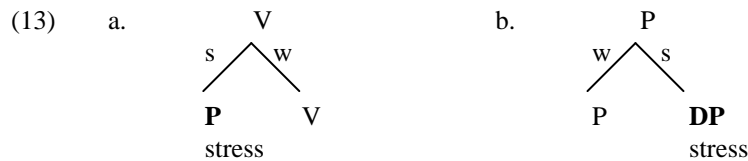
When the combination [Prt + V] is again labeled V (V projects), it is plausible that a DP complement of the complex V is most likely to appear with an accusative marking. The complement of the Prep, by contrast, may be in for an oblique marking. In this way, the differences between Prep and Prt, in *order*, *stress* and *selection*, as listed in (10), fit in with the category assignment rules in (11) and (12).

5. The difficulty

The argument frame [Prep __]$_{PrepP}$ from (11) fits the verb *zitten* ('sit'). There is an oblique prepositional object as in (14a). The pronominal variant (13b) is an oblique *er/d'r*. It shifts out of the predicate, stranding the preposition in (14b).

(14) a. ga   maar niet   [op de stoel] zitten
        go   but   not   on the chair sit
        'you may better not sit on the chair'
    b. ga   er$_i$/dr$_i$ maar niet [op t$_i$] zitten
        go   there   but not   on t$_i$ sit
        'you may better not sit on it'

Here we predict successfully a complication. When the provisional category assignments from (11)-(12) are still active, the stranded Prep in (14b) cannot as such be recognized, especially because the stress sometimes falls on the stranded Prep when the DP pronominalizes as *er/d'r*. The stranded stress bearing Prep will then rather be taken together with the adjacent V and get analyzed as a [Prt + V]. The oblique object *er/d'r* cannot yet be analyzed as a shifted Prep complement. There is in the simplified view of the language acquisition procedure no Prep unless there is support by an explicit DP. We propose to reconstruct this as follows. The acquisition procedure will perform a double categorization. It will assign a Prep in (15a) and a (pseudo) Prt in the stranded variant (15b).

(15) a. ik ga   niet op$_{<+D>}$ die stoel   *zitten*
        I go   not   on that chair   sit
        'I will not sit on that chair'
    b. ik ga (er/d'r) niet op$_{<Pred>}$- *zitten*
        I go (there)   not on-sit
        'I will not sit on it'

There is evidence from spontaneous production data that this is indeed what is going on in child Dutch. The shifted object pronouns (oblique and accusative) are left out in half of the relevant cases roughly between the age of 2-3. This may indicate that, being grammatical elements, the pronominal objects are not yet easy to interpret. See the percentages for five Dutch children in CHILDES (MacWhinney, 2006) in (16). The children are from the Groningen corpus (Abel, Daan, Iris and Matthijs) and from the Van Kampen corpus (Sarah). The target language includes the object pronoun between parenthesis (*het*, *er*). The figures give the percentage of target failure for [Prt-V] and [Prep-V]. The [Prep-V] is reanalyzed as a pseudo [Prt-V].

(16) Percentage of 'object-drop' in local (binary) [P + V] context

	(*het*) [Prt-V] frame kwil (*het*) **op-tillen** wanna (it) up-lift 'wanna lift (it) up'	(*er*) [Prep-V] frame → [Prt-V] kwil (*er*) **op-zitten** wanna (there) on-sit 'wanna sit on (it)'
Abel 2;0-3;1	48% (n 25/52)	30% (n 20/67)
Daan 2;2-3;1	54% (n 44/82)	52% (n 50/97)
Iris 2;8-3;5	42% (n 19/45)	42% (n 41/97)
Matthijs 2;2-3;0	53% (n 56/105)	54% (n 50/93)
Sarah 1;8-2;9	58% (n 46/79)	49% (n 30/61)

The counts in (16) show how the accusative pronoun *het/'t* ('it'), or another weak accusative pronoun like *'m* ('him'), *d'r* ('her'), is not realized in case of transitive constructions based on a [Prt-V] frame. For example, the [Prt-V] *op-tillen* ('lift up') lacks an object in roughly 50% of the cases. This is mainly due to not realizing the pronominal object, see (17c). When the object is present in early child language (the other 50%), it tends to be a DP/noun or a demonstrative (*dat/die* 'that', *dit/deze* 'this') (Van Kampen 2004), as in (17a,b).

(17) a. ik ga   (het) steentje optillen
      I go   (the) stone   up-lift
      (I am going to lift (the) stone up)
    b. ik ga   dat optillen
      I go   that up-lift
      (I am going to lift that one up)
    c. ik ga   (het)   optillen   (roughly 50%, see (16))
      I go   (it)   up-lift
      (I am going to lift (it) up)

The bracketing of the pronominal object (*het*) in (17c) indicates that the pronoun is not realized.

The pseudo Prt construction shows the same type of target failure. The oblique object is realized as a DP/noun (18a). It is also possible that the strong oblique locative adverbs (*daar* 'there', *hier* 'here') appear, as in (18b). These adverbs have already been acquired as locative adverbs. Although they also function (in the adult grammar) as strong oblique pronouns related to the stranded Prep, such an interpretation is not an option before the P-element itself has been identified as a Prep that requires an oblique object. Hence, we feel entitled to leave the locative adverbs (18b) out as an acquisition context for stranded Prep. Non-locative *daar* is not attested before the third birthday and the current grammar can and hence will get away with (18b) as a free local adverb. The weak oblique pronoun *er/d'r* (18c) is marginal and non-presence of the oblique object (18d) is the rule.

			Abel	Daan	Iris	Matthijs	Sarah
(18) a.	ik wil	op (de) stoel zitten	60%	28%	53%	35%	43%
	I want	on (the) chair sit					
b.	ik wil	daar op zitten	4%	17%	1%	7%	3%
	I want	there on sit					
c.	ik wil	er op zitten	6%	3%	4%	4%	5%
	I want	there on sit					
d.	ik wil	(er) op zitten	30%	52%	42%	54%	49%
	I want	(there) on sit					

The percentages in (16)/(18) show that the weak pronouns are for some time insufficient guides to get the difference between Prep verbs and Prt verbs. Hence, one of the reviewers suggested that (16) demonstrates only how "children have similar developmental issues with accusative and oblique object pronouns". This seems a valid perspective, but there may be a better way to demonstrate a target failure

such that the stranded Prep is bound to be analyzed as pseudo-Prt. Let us oppose only the constructions with a real weak pronoun *er/dr* (18c) versus the construction with a **'gap'** (18d). See (19). Now we find a much higher target failure for uncontroversial oblique objects and hence a much stronger argument that stranded [Prep +V] was bound to be interpreted for some time as pseudo [Prt + V] in child Dutch according to the acquisition context in (12).[2]

(19) Presence versus Absence of *er/d'r* in [P-V] constructions

	Presence of *er/dr* Type (18c) 'kwil *er/dr* op-zitten'		Absence of *er/dr* Type (18d) 'kwil op-zitten'	
Abel 2;0-3;1	17%	(n 4/20)	**83**%	(n 20/24)
Daan 2;2-3;1	6%	(n 3/53)	**94**%	(n 50/53)
Iris 2;8-3;5	2%	(n 1/42)	**98**%	(n 41/42)
Matthijs 2;2-3;0	7%	(n 4/54)	**93**%	(n 50/54)
Sarah 1;8-2;9	9%	(n 3/33)	**91**%	(n 30/33)

To summarize, our present concern is not so much the omission of weak pronouns in child Dutch as such, but the fact that this inevitably leads to a categorization error in the case of Prep-stranding. The acquisition procedure cannot but interpret *opzitten* in (15b)/(18d) as a particle verb [Prt-V]. The percentages of *er*-drop in (19) support the view that reduced pronouns (*er/d'r* in this case), largely fail to be noticed by the child and therefore cannot be interpreted as complement of P.

Additional evidence comes from the form of the pronoun when it is present, mainly after the third birthday. The shifted oblique *r*-pronouns are often replaced by their accusative counterparts, although there is no support from the input for this replacement. In adult Dutch the oblique *r*-pronoun is obligatory when the Prep is stranded. The accusative pronoun in child Dutch is compatible with a (pseudo) Prt analysis and not with a [Prep t$_{DP}$] analysis. A particle verb may very well realize its object argument by means of an accusative configuration as (20b) exemplifies. Incorrect accusatives as in (20a), where the stranded Prep is overgeneralized as a Prt appear quite frequently in child Dutch between the age of 3-6 (Van Kampen, 1996; Coopmans & Schippers, 2008). Some examples are given in (21).[3]

(20) a. *ik ga die (stoel) niet *op$_{<Prt>}$- zitten*
  I go that (chair) not on-sit
  (I won't sit on it)

  b. ik ga die (stoel) niet *op$_{<Prt>}$- tillen*
  I go that (chair) not up-lift
  (I won't lift it up)

(21) a. *die* kunnen we mee ballen   (Matthijs 2;06)
  that can we with ball
  'we can play football with that'

  b. *dat* kan ik niet mee dansen   (Daan 3;01)
  that can I not with dance

---

[2] Admittedly, some children may sometimes use *er/d'r* before the age of 3. However, it is questionable whether they interpret *er/d'r* already as a complement of P. Abel, for instance, has more instances of *er/d'r* than the other children, especially in the context without a non-finite V (*ik wil niet erop* 'I want not thereon'). However, at some point he starts doubling *er/d'r* (*ik wil er niet erop* 'I want there not thereon'). We counted 24 of such doublings after 3;0. Therefore, we assume that initially Abel does not analyze *erop* as [*er* + P].

[3] One of the reviewers suggested that the data in (21) could in principle be interpreted as "truncated dislocation constructions, containing an oblique (but truncated) pronoun", i.e. (21a) *Die, (daar) kunnen we mee ballen* ('That, (there) can we with ball'). This would, though, shift the problem to the unattractive question why the non-truncated forms (**Die, daar kunnen we mee ballen*) never show up in child Dutch and even are ungrammatical in adult Dutch (Barbiers et al., 2005, p. 73).

'I cannot dance with that'
- c. weet je   *wat*   ik   over heb gedroomd?   (Sarah 3;11)
     know you   what   I   about have dreamt?
     'do you know what I have dreamt about?'
- d. *die* heb ik   in gekleurd   (Sarah 2;9.7)
     that have I   in colored
     'I have colored therein'
- e. *dat elastiek*   heb ik   mijn voeten   op gezet   (Sarah 4;9.13)
     that elastic   have I   my feet   on put
     'I have put my feet on that elastic'
- f. ik weet   hoeveel   we *mee* zijn   (Laura 6;10)
     I know   how many   we with are
     'I know with how many we are'

The actual stress pattern (from the recordings) might show that we are on the right track. For instance, Sarah said *die heb ik **in** gekleurd* (21d) with a clear stress on *in*. The sentence was uttered in the context where she has scratched in a book with her pencils. The adult stress obligatorily falls away from the Prep on the V *kleuren* ('to color'): *daar heb ik in **gekleurd***.

The reverse, [Prt + V] taking an oblique object, is not attested. The present approach fits in with this asymmetry. It predicts a misanalysis (an acquisition failure) of stranded Prep as Prt, but it leaves no room for a misanalysis of Prt as Prep. One needs extended corpora up to year six to find the systematic appearance and disappearance of these Prep-stranding errors.

The long acquisition period for the shifted oblique pronoun (between the third and the sixth birthday) can be interpreted as a slow retreat from the (pseudo) [Prt + V] analysis. The [Prep + DP] allows a pronominalization [Prep + t_{er}] + V. This pronominalization is present in a local binary frame [er + Prep t_{er}] when no intervenient material is between Prep and *er* and *er* is minimally shifted to the left, see (22).

(22) Local (binary) context [er + P] for shifted *r*-pronouns
     ik   doe mijn schoen   er-in, er-op, er-uit
     I   put my shoe   there-in, there-up, there-out

The locally cliticized pronominal *er*-P/*d'r*-P is abundantly supported by input, and will give rise to the interpretation [*er/d'r* [Prep + t_{DP}]]. When as a first step *er*-P is interpreted as a pronominal variant of [Prep + DP], the acquisition procedure will reach a second step, the obligatory shift of the weak *r*-pronoun beyond the predicate. The pseudo [Prt + V] analysis does not have this implication and it will be phased out to the extent that the shifted oblique pronoun *er/d'r* is ruled in. Retreat from (pseudo) Prt to stranded [Prep + t_{DP}] is slow because it operates as a blocking procedure for the (pseudo) Prt and must take place for each lexical item separately.

6. Conclusion

The background of the present note is our general conviction that the language acquisition procedure contains a sub-procedure for identifying the syntactic category of each lexical item. The sub-procedure should work item by item. An item that got its category identified is stored in the lexicon with that category. The category assignment rules are input-oriented acquisition procedures. They are based on the minimal (binary) frames that appear due to the input reduction of early child language. These language-specific frames offer an alternative to an acquisition procedure based on parameter-setting frames as a priori options of the mind (Van Kampen 2009).

The acquisition procedure applies at first a massive input reduction. The input reduction is based on ignorance about the target system. It is a decoding procedure that temporarily leaves out any element that cannot yet be interpreted in a binary frame. The sub-procedure for category assignment crucially operates within that simplified medium. The acquisition graphs for systematic predicate marking by I^o and for reference marking by D^o demonstrate that the categories V(erb) and N(oun) get identified

roughly before the third birthday. The definition of verbs by means of tense marking corresponds with the traditional view on category definition. The same holds for the definition of nouns by means of articles or case inflection.

The context factors V^o and D^o allow the category assignment rules for Prep (rule (11)) and for Prt (rule (12)) in Dutch. The fact that the same items may appear as Prt or as Prep (given the appropriate verb) requires an historical explanation. Particle verbs are productive and new particle verbs may arise. New particle verbs may historically arise from the reanalysis of an element left-adjacent to the verb, like a postposition (see the overview in Blom 2005 and references cited therein). The present cases in child language suggest that particle verbs may arise as well from prepositions that appear V-adjacent due to Prep-stranding. Stowell's (1981, 448) conjecture that Prep-stranding appears only in languages that have particles may need to be reversed. As a matter of fact, most Prt elements function as well as Prep. Our present suggestion for that coincidence is that P-stranding in OV languages has fed the amount of particle verbs in child language as well as in the history of language.

The central point is that the simplified binary context frame for Prt acquisition in (12) predicts an actual overgeneralization. The stranded Prep has to be analyzed in early child Dutch as (pseudo) Prt. It was observed that the child leaves out most of the weak pronouns between the second and third birthday. Stranded prepositions in the adult input cannot be interpreted as $[\text{Prep } t_D]_{\text{PrepP}}$ at this stage. Moreover, the pseudo-Prt acquisition failure causes another acquisition error after the third birthday. Often accusative pronouns pop up rather than the adult oblique pronouns, whereas the reverse cases were not attested. The faulty accusative phenomenon supports the idea that the stranded Prep may still be interpreted as Prt. Retreat from these acquisition errors starts once oblique *r*-pronoun shift becomes an option. Yet, the retreat is slow, since this operates as a blocking procedure for the pseudo Prt. It must take place for each lexical item separately.

References

Barbiers, Sjef, Hans Bennis, Gunther De Vogelaer, Magda Devos & Margreet van der Ham (2005) *SAND. Syntactische atlas van de Nederlandse dialecten. Deel 1: Pronomina, congruentie en vooropplaatsing.* Amsterdam: Amsterdam University Press. http://dare.uva.nl/aup/en/record/173114

Besten, Hans den. (1983). On the interaction of root transformations and lexical deletive rules. In Werner Abraham (Ed.) *On the formal syntax of westgermania* (pp. 47-131). Amsterdam: John Benjamins.

Blom, Corrien. (2005). *Complex predicates in Dutch. Synchrony and diachrony.* PhD dissertation VU (LOT)

Cartwright, Timothy & Michael Brent. (1997). Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition, 63* (2), 121-170.

Coopmans, Peter & Rianne Schippers. (2008). Preposition stranding in development. In Marjo van Koppen & Bert Botma (Eds.) *Linguistics in the Netherlands 2008* (pp. 75-86). Amsterdam: John Benjamins.

Emonds, Joseph. (1985). *A unified theory of syntactic categories.* Dordrecht: Foris.

Evers, Arnold & Jacqueline van Kampen. (2008). Parameter setting and input reduction. In Theresa Biberauer (Ed.) *The structure of parametric variation* (pp. 483-515). Amsterdam: John Benjamins.

Jakobson, Ramon. (1942). *Kindersprache, Aphasie und allgemeine Lautgesätze.* Uppsala: Uppsala Universitets årsskrift 9.

Kampen, Jacqueline van. (1996). PF/LF convergence in acquisition. In K. Kusumoto (Ed.) *Proceedings of the North East Linguistic Society* (NELS) *26* (pp. 149-163). University of Mass., Amherst: GLSA.

Kampen, Jacqueline van (1997) *First steps in wh-movement.* Delft: Eburon.

Kampen, Jacqueline van. (2004). Learnability order in the French pronominal system. In Reineke Bok-Bennema et al. (Eds.) *Selected papers from Going Romance 2002* (pp. 163-183). Amsterdam: John Benjamins.

Kampen, Jacqueline van. (2009). The non-biological evolution of grammar: Wh-question formation in Germanic. *Biolinguistics, 2-3* (3), 154-185. http://www.biolinguistics.eu/index.php/biolinguistics/article/view/83/111

Lebeaux, David. (2000). *Language acquisition and the form of grammar.* Amsterdam: John Benjamins

MacWhinney, Brian. (2006). *The CHILDES Project: Tools for Analyzing Talk*, 3rd edition. Mahwah NJ: Lawrence Erlbaum Associates.

Maratsos, Michael & Mary Chalkley. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K.E. Nelson (Ed.) *Children's language* (Vol. 2, pp. 127-214). New York: Gardner.

Mintz, Toben. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition, 90* (1), 91-117.

Morgenstern, Aliyah & Martine Sekali. (2009). What can child language tell us about prepositions? A corpus-based study of cognitive and social-pragmatic factors. In Jordan Zlatev et al. (Eds.) *Studies in language and cognition* (pp. 261-275). Cambridge: Cambridge Scholars Publishing.

Neeleman, Ad & Fred Weerman. (1993). The balance between syntax and morphology: Dutch particles and resultatives. *Natural Language and Linguistic Theory*, *11*, 433-475.
http://www.springerlink.com/content/r45r670732ww772r/

Pinker, Stephen. (1984). *Language learnability and language development*. Cambridge MA/London: Harvard University Press.

Redington, Martin, Nick Chater & Steven Finch. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22* (4), 425-469.

Riemsdijk, Henk van. (1973). The Dutch P-soup. Unpublished manuscript University of Amsterdam

Riemsdijk, Henk van. (1978). *A case study in syntactic markedness: The binding nature of prepositional phrases*. Dordrecht: Foris.

Ruys, Eddy. (2008). Stranding, weak pronouns, and the fine structure of the Dutch Mittelfeld. *Natural Language and Linguistic Theory*, *26*, 547-587. http://www.springerlink.com/content/ev51746668784174/

Schippers, Rianne. (to appear) The acquisition of particle verbs in Dutch. PhD dissertation Utrecht University.

Snyder, William (2007) *Child language*. Oxford: Oxford University Press.

Stowell, Tim. (1981). Origins of phrase structure. PhD. dissertation MIT Cambridge Massachusetts.

Tomasello, Michael. (1987). Learning to use prepositions: a case study. *Journal of Child Language*, *14*, 79-98.