

Bootstraps at two for lexicon and discourse
Appeared in Proceedings of ELA (2001), submitted to L.A., December 2001

Jacqueline van Kampen
UiL OTS, Utrecht University
jacqueline.vankampen@let.uu.nl

This article presents the acquisition graphs for D/I-marking in early Dutch child language. Two major considerations will be derived. Firstly, the acquisition of the D/I-system will be construed as the key to the N/V-category assignment, a reversal of Pinker's bootstrapping scenario (Van Kampen 1997). Secondly, empirical facts will support the idea that D/I-marking is the universal step for setting up discourse grammar.

Data from early child language show that 1) content words start as category-neutral signs X, and not as N or V; 2) functional words are ideal bootstraps for category assignment, because of their high token frequency. Curiously enough, the functional words themselves do not start as D^o or I^o. They rather appear as illocution operators added to a content sign. E.g. in Dutch child language, later demonstratives appear at first as illocution markers of gesture deixis in presentational utterances. Later finite verbs start as illocution markers of subjective modality in deontic/volitional utterances. These early illocution signs are completely situation oriented. Only later they are interpreted as pronominal markers of referential opposition (D^o) or syntactic predicate markers (I^o). As signs of syntactic deixis, D^o/I^o identify distinctions that maintain a presence of information the speaker may refer back to by the grammatical devices of discourse grammar. Empirical support for this view follows from the acquisition of discourse anaphora and discourse connectives.

1 Pinker's bootstrapping scenario

It would be hard to assign much grammatical structure without different categories from which grammatical structure can be projected. Since functional categories are typically left out in early child language, it seems a possibility to start the development of a categorial system with the universal categories N and V. The more language specific design of the D^o and I^o categories might then be detected as a frequent context of the earlier universal categories N and V. Pinker's (1984) cognitive bootstrapping mechanism follows that path. It derives content categories from general notions, like 'thing' (N) or 'action' (V). Presumably, an innate set of procedures would subsequently map the primary syntactic categories (N and V) onto the phrase structure ($S \rightarrow NP VP$). The child's early word combinations are in Pinker's theory assumed to project the full category labels NP and VP. I suppose that, within such a view, the lexical conceptual structures (NP/VP) will in turn constitute a platform for learning the functional categories I^o, D^o, C^o and their projections (DP/IP/CP). In short, the acquisition procedure would translate a conceptual cognitive system (thing/action) into a system of lexical categories (N/V) and these would open the way for functional categories (D/I). The translation from the cognitive orientation towards the grammatical point of view was called 'bootstrapping'. Pinker revised his semantic bootstrapping theory thoroughly in Pinker (1987,1989). Nevertheless, he continued to defend the position that lexical categories are accessible to the child in a syntax-free manner.

Gleitman (1990) used the same metaphor 'bootstrapping' as well, but in a more loose way. Whereas Pinker reserved the term 'bootstrapping' for a learning procedure that switches from general cognition to grammar, Gleitman applies 'bootstrapping' to the support one component of the system may give to another part of the grammar (see also Pinker 1994). The acquisition of the lexicon, she argues, is supported by the emergence of

the parts of speech system into N/V/A/P and by the emergence of theta frames.¹ A grammatical frame would strongly support the development of a cognitive frame. For instance, the cognitive content of a verb is 'syntactically bootstrapped' through the child's attention to the argument structure, cf. also Hale & Keyser (1993) for theta frames that are constrained by a prior X-bar system.

Pinker's semantic bootstrapping towards lexical categories has certain drawbacks. It must begin with excluding the existence of action nouns and stative verbs as a matter of principle. The pre-grammatical conceptual orientation can only be meant as a temporary provisional start of the category-learning procedure. Thereafter, something must cause a dramatic shift towards formal grammatical categories. Such a shift towards a formal criterion is also needed for conversion types that are quite common in child language (*sleep, shower, play, walk*). Consider for instance the following. It is not possible to determine the <+/-N> and <+/-V> status of items like *sleep, breath, shower, play, walk, laugh, stay* etc., if there is no context like, for example, *a sleep* or *I sleep*. This holds for the adult language in these specific cases as well as for child language in general. The category status of e.g. *sleep* is determined by its syntactic context. Categories that are bound to be in one of these contexts are N or V: **I tree / a tree: tree* <+N> and *I think / *a think: think* <+V>. A second drawback is that the notion 'thing' has to be stretched up to contain spaces (*room, cellar, garage*), fluids (*milk, water, juice*), states (*rain, weather, dark, health, thirst, hunger*), and that the notion 'action' will begin to cover all type of occurrences and properties (*falls, glimmers, burns, rains*).² The young learner needs some flexibility, but if that flexibility is readily available, the distinction between Ns and Vs becomes uncertain and some help from syntactic context may be welcome. Here we are back to the original problem. There is no syntax if there are no categories and there can be no categories if there is no syntax. There is a hole in your bucket, dear linguist, dear linguist.

2 The pragmatic scenario

Suppose now that the early content words during the single-word and two-word phase all belong to a neutral category X and that the characterization of these as <+N> or <+V> is due to an over-interpretation by the adult. A single sign X, syntax free and therefore category neutral, can be used, and is used by the child, as a name X<+ref>, as in (1)a, or as a characterization X<+pred>, as in (1)b. This depends on the pragmatic intentions of the speaker, not on any property already inherent in the child's understanding of the sign. See also Ketrez & Koç (2000), who considered category neutrality in Turkish child language. The point was already made in Lyons (1977:649, 1979:90). Van Kampen (1997:chap2,3) takes up the same idea and extends it into a scenario for the acquisition of whatever syntactic categories. The examples below are given for Dutch and English child language,

¹ Gleitman (1990) was well aware that she did not address Pinker's core problem how to turn a non-grammatical interpretation into a grammatical point of view (<+/-N> <+/-V>). See her footnote 9, page 27: "Pinker (1984) actually reserved the term semantic bootstrapping for machinery that assigns words to lexical categories. For expository convenience, however, I take the liberty of using his expression to refer to his proposals at their broadest for extracting verb meanings from extra-linguistic context." Hence, Gleitman does not take her proposal to be an alternative answer to Pinker's original problem. My proposal, by contrast, is meant as an alternative to Pinker.

² A similar problem arises for the acquisition of active and passive predicates (see Pinker 1989: 413; Gropen et. al. 1991). Subject-predicate relations vary from an inherent actor/action relation to theme/change-of-state relations. If the subject-predicate relation is at first identified by the child as an actor/action relation, instead of the other way around, it must be that the appearance of theme/change-of-state predications implies a reversal of the strategy. And yet, passive and change of state predications appear as early as predicate I-marking with and without past participle marking (Verrrips 1996).

but the procedure is intended as a general and crucial operation by the human language acquisition device.

(1) *Pragmatic switch of intention*

a. <+ref>

|
X

beertje (bear)
(proper name with fixed reference,
i.e. name for the child's cuddly bear)

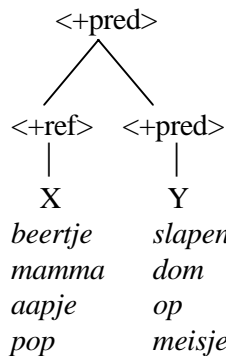
b. <+pred>

|
X

beertje (bear)
(characterization as it might hold
for a set of animals)

The referential versus characterizing use of the neutral content sign X extends into the two-word utterances. Consider the following examples of utterances with two content elements, both without a category label. The features <+ref> and <+pred> indicate the pragmatic intention of the sign.³

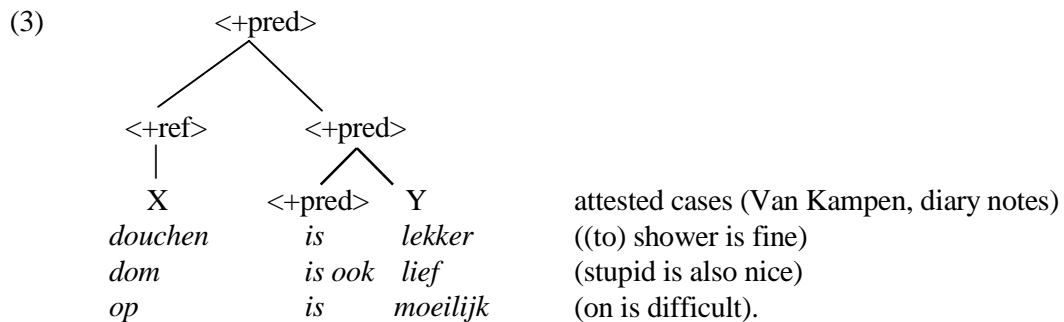
(2)



examples from Sarah (Van Kampen, CHILDES)
(bear sleep)
(mummy stupid)
(monkey on)
(dolly girl)

The whole construct stands for the situation. The characterizing signs are applied to the reference intended sign X. This is the invention of predication. The constructs in (2) do have a grammatical shape. There is order (subject-X first) and stress (on predicate-Y) (Van Kampen 1997). Yet, it is not necessary to assume that the child already needs labels for the categories N or V in (2) and could not proceed by means of category-neutral content signs. All category-neutral signs may appear in the characterizing predicate position on the right, but quite soon as well in the referential subject position on the left, and one gets *douchen is lekker* ('(to) shower is fine'), *dom is ook lief* ('stupid is also nice'), *op is moeilijk* ('on is difficult').

³ The use of <+pred> for single word utterances that 'characterize' rather than 'name' is not completely fortunate, as Sergio Baauw pointed out to me. Type theory at least requires two signs for an predication <e, <e,t>>, a name entity <e> and a predicate <e,t>. A single sign cannot carry a predicate's truth-value. Maybe a better formulation for the signs in (1) is given in Lyons (1979:90). He distinguished the naming versus the characterizing function of single words utterances as proto-reference versus proto-predication.



Besides two-word predicates of type (2), constructs of type (4) below are very frequent in early child language (see Braine 1963). They contain a category-neutral content sign (word or phrase) combined with a fixed functional non-content sign that marks the pragmatic use of the illocutionary act as a unit in a language game (see for various analyses of the use of non-content signs in early child language Brown 1973, De Haan 1987, Van Kampen 1987/1994/1997, Lebeaux 1988, Hoekstra & Jordens 1994, Roeper 1996, Powers 1996a/1996b, Powers & Lebeaux 1998, Penner, Tracy & Weissenborn 2000, Hulk 2001, and others).⁴

The pragmatic constants will later develop into a variety of grammatical devices, but their initial illocution operator function merely characterizes a presentational, volitional or deontic predicate or a turn in the conversational interaction. These early illocution operators are completely non-discourse bound. Only later they are interpreted as discourse bound markers of referential opposition (D⁰) or syntactic predication (I⁰). The examples in (4) are taken from the files of Sarah (Van Kampen corpus, CHILDES). Table (4)a lists the major operator+X constructs in Dutch child language, and (4)b the major X+operator constructs, where X stands for any content sign, i.e. a single word or a phrase.⁵ The elements in brackets are optionally combined with the illocution operator. The order operator+X or X+operator is rather strict.⁶ Some positional inconsistencies may occur, but they stay below a negligible 10% (for such a point of view, see Evers & Van Kampen 2001).

⁴ The use of a category X is not uncommon as an expository notation. E.g. the operators *want* {P:*up*; V:*see*; N:*car*; A:*tight*} and *more/no* {P:*up*; V:*read*; N:*ball*; A:*dirty*} in child language are characterized as *want X* and *more/no X* in Powers (1996a) and Powers & Lebeaux (1998). It is proposed here to take the notation seriously, as in Lyons (1977, 1979), and to derive the categories <+/-N>, <+/-V> later on from discriminating functional contexts.

⁵ Lyons (1977) discusses two-word sequences in English child language with a deictic operator (D). The pattern may be D+X or X+D. According to Lyons, the child makes no distinction between 'names, common nouns and verbs' yet (p. 649), i.e. X is category-neutral.

⁶ I do not define the notion 'operator' here, but see for instance Heim (1982:143) for the operator status of negation, temporal adverbs and modals. The spell-out of information structure (separation of scope-assigning operators and theta structure) in child language was advanced in Van Kampen (1997).

(4) a.

illocution operator (+ category-neutral X)	illocution operator (+ category-neutral X)	illocution operator (+ category-neutral X)
<i>deictic operator</i> dat (is) X dit (is) X die (is) X deze (is) X hier (is) X daar (is) X (dat) niet (is) X	<i>modal operator</i> (ik) kan (wel/niet) X (ik) mag (wel/niet) X (ik) hoef (wel/niet) X (ik) wil (wel/niet) X ikke/kwi/ ik ook X ik niet X	<i>turn taker</i> nou X nog X ook X even X zo X niet X

dat/dit/die/deze/hier/daar is X (that/this/here/there is X); *niet is* X (not is X); *ik kan wel/niet* X (I can indeed/not X); *ik mag wel/niet* X (I may indeed/not X); *ik hoef wel/niet* X (I need indeed/not X); *ik wil wel/niet* X (I want indeed/not X); *ikke/kwi/ ik ook* X (I sjwa/wanna/I also X); *(ik) niet* X ((I) not X); *nou/nog/ook/even/zo/niet* X (now/yes (=more)/also/just/so/not X)

b.

(category-neutral X +)
illocution operator
<i>place pointer</i>
X nou ? (X then?)
X daar (X there)
X hier (X here)
X ook (X also)
X niet (X not)

Some examples of type (4)a and (4)b are given in (5)a and (5)b. For a more complete overview, see Appendix A.

(5) Dutch (Sarah, Van Kampen corpus)⁷

a. niet∂ bad zwemmen	(1;11.15, week 102)	(not∂ (is) (to) ‘bath swim’)
(i)kwi(l) uit	(1;11.15, week 102)	(wanna out)
dit∂ lachen	(2;0.17, week 107)	(this∂ (is) (to) laugh)
deze niet leuk	(2;0.17, week 107)	(this (is) not nice)
mag wel kleuren	(2;1.10, week 110)	((I) may indeed color)
ik ook kleuren	(2;1.10, week 110)	(I also (want to) color)
niet een centje hebben, hoor	(2;1.10, week 110)	((this is) not ‘have a cent’)
nog paardje ⁸	(2;1.10, week 110)	(‘more’ (you play) horse)
deze ook nog (ko)nijnen	(2;4.9, week 123)	(this (are) also yet rabbits)
deze niet nou vogel	(2;4.2, week 122)	(this (is) not now bird)
nou donker	(2;4.27, week 124)	(now (you make) dark)
kan deur	(2;5.22, week 129)	((I) can (open the) door)
dan even klaar	(2;6.28, week 133)	(then just ready)
dat niet een rails	(2;7.11, week 137)	(that (is) not a rails)
b. zitten ook	(1;9.10, week 93)	((to) sit also)
is jou niet	(2;3.16, week 120)	(is not you(rs))
clown nou?	(2;4.27, week 124)	((where is) clown now?)
die stuk daar	(2;7.16, week 137)	(that piece (is/must) there)

The closed class illocution operators have each a much higher frequency (> 10x) than any of the open class content labels. The number of occurrences in Sarah’s ≥ two-word

⁷ The symbol ∂ stands for a ‘schwa’.

⁸ *nog* is used by the Dutch child in the context where the English child uses ‘more’.

utterances are given in (6)a and (6)b. Proper names, although clearly referential, have not been counted as content signs and are left out of the comparison (e.g. *mamma* ('mummy'), *pappa* ('daddy')). The content signs X will appear later in the adult grammar as N, V, P and A. During the first period (till about week 120) most of the elements in bold characters listed in (4)a,b were used as pragmatic constants, see the quantifications in Appendix B.

- (6) Number of occurrences for single items⁹ (Sarah, week 86-150)
- a. *Illocution operators*: between 50-300
 - b. *Content elements X*: between 1-30

The early use of D-elements (*dit/dat/die/deze*) and I-elements (*is/wil/mag/kan/moet*) may signify no more than mere utterance announcers in early child language, i.e. signs that announce a presentational utterance, the performing of an act, or a turn in the conversation, see (7).¹⁰ They are stereotypes that mark the category-neutral X as the illocutionary aim of a pointing gesture, a wish, an intention, a question, or a refusal. One might say that these illocution operators meta-linguistically introduce the utterance with colons.¹¹

It seems a mistake, intuitively speaking, to characterize these initial utterance announcers syntactically as demonstratives, modal verbs, sentence adverbs, subject clitics etc. from the start, since the two-year old speaker need not distinguish between them yet. The pragmatic intention of the constants may nevertheless have been clearly and correctly assessed.

- (7) *Illocution operators*
- | | | |
|------------------------|--------------|-------------------------------|
| a. for presentationals | dat: beer | (that: bear, cf. Braine 1963) |
| b. for wish/comment | ikð: beer | (Ið: beer) |
| c. for wish | (k)wil: beer | (wanna: bear) |
| d. for command | moet: beer | (must: bear) |
| e. for assertion | isð: beer | (isð: bear) |

The illocution operators develop into a more refined set of syntactic categories when the pragmatic intentions become more implicit as grammatical structure differentiates (Van Kampen 1987/1997, Roeper 1996). The D-elements develop into demonstratives (use in object position, without intonation break, alternation with articles and possessives), and the I-elements are used in I⁰ with other auxiliaries and finite verbs.¹² The representation of

⁹ The only exception *boek-je* ('book-let', 65 times) is due to a favorite contact form of the conversation on the tapes.

¹⁰ The various sentence adverbs (*nou/nog/ook/even/zo/dus* 'now/yes/also/just/so') are quite frequent and subtle means in adult Dutch. They are reduced to a few pragmatic functions in early child language: presentational gesture, emphasis, turn taker in a conversational exchange (Van Kampen 1997:79f, 2000b). For alternative views on sentence adverbials in Dutch/German see Powers (1996b), Penner, Tracy & Weissenborn (2000). I have labeled the sentence adverbs intuitively as functioning as turn-takers or place-pointers in early child language.

¹¹ See Radford (1990:75f) for the meta-linguistic use of *that/this*, Drozd (1995) for the meta-linguistic use of *no*, Van Kampen (1987), Hoekstra & Jordens (1994) for the meta-linguistic use of *niet*, and Powers & Lebeaux (1998) for the meta-linguistic use of *more* and *no* in child language.

¹² In a typological study of predicate licensing in 410 languages, Stassen (1997) observes that in many languages the morphological form of the copula reminds of the morphological form of subject pronouns. He shows how copula in presentationals may have a historical origin as a pronoun (Stassen 1997:99). This seems less strange if we see how both elements are likely to be picked up in child language as utterance announcers, see real examples of the copula as an utterance announcer in (i)-(iii). The Dutch and English examples appeared mainly at the moment that the referential *dat/that* started to be used outside of the

these markers as pronouns and modals later on, does not contradict their initially merely pragmatic function in linguistically standardized interactions. The development of contexts for the categories N and V seems more perspicuous, as we will see in the next two sections.¹³

3 Bootstrapping N

Consider again the constructs in (7) with an illocution operator and a content element. The deictic demonstratives *dat/dit/die/deze/hier* ('that/this/here') for the pointing gesture in presentationals have at first an intonation break (see for presentationals in child language Van Kampen 1997:p.23,127).¹⁴ They cannot be used in a predicative way and are exclusively referential. In 'picture looking' games, or somewhat freer 'naming games' they may refer to a thing or person, but also function as a utterance announcers for an entire situation as in the examples in (8). Identical patterns with the deictic operator referring to an entire situation, are probably present in English, French and German child language as well, see Appendix C.¹⁵

presentational construction, week 120 for Sarah (see (15)b) and week 112 for Nina (Suppes corpus. See Van Kampen 2001 for Nina's acquisition graph of *be*).

- | | | | |
|-------|----------------|-----------------------------------|------------------------------|
| (i) | <u>Dutch</u> | (Sarah: week 120) | |
| | a. | is gieter/bordje/water/plakkertje | (is can/plate/water/sticker) |
| | b. | is heet/lekker/vies | (is hot/nice/dirty) |
| | c. | is mij/jou niet | (is mi(ne)/not you(rs)) |
| | d. | is op | (is on) |
| | e. | is koud buiten | (is cold outside) |
| | f. | is deze koud buiten | (is this one cold outside) |
| (ii) | <u>English</u> | (Nina: week 112) | |
| | a. | is Mommy living room | |
| | b. | is Mommy's living room | |
| | c. | is a girl | |
| (iii) | <u>French</u> | (Grégoire 1;10.20 week 98) | |
| | a. | est crocоди(le) | (is crocodile) |
| | b. | est casquette d' Adrien | (is cap of Adrien) |
| | c. | est chaussette Victor | (is sock Victor) |

If the adult language is on the border of pro-drop or copula-drop, a reinterpretation of the pragmatic constant as I^o or D^o becomes possible.

¹³ The initial meta-linguistic sentence adverbs/negation develop into IP/VP adjuncts or modal verbs later on.

¹⁴ In adult Dutch *deze/dit* are the proximal demonstratives and *die/dat* the distal demonstratives; *dit* and *dat* are used with neuter nouns. Early child language does not make these distinctions yet, see section 6.

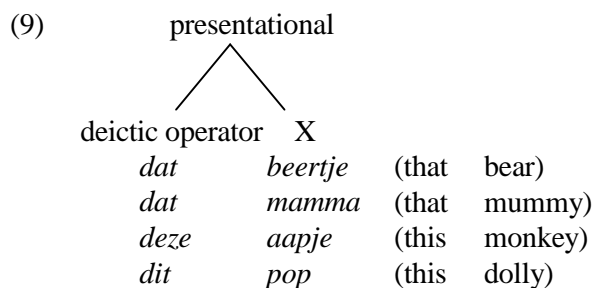
¹⁵ Modal operators, turn takers and negation operators may also metalinguistically refer to an entire situation (cf. Van Kampen 1987, Drozd 1992, Radford 1996, Roeper 1996, Powers & Lebeaux 1998). See also Appendix C. Much unlike the present category-neutral analysis, Powers & Lebeaux (1998:49,55,70) assume that the Event complement of the *more/no* operator already realizes an IP.

(8) *Deictic operators in presentationals*

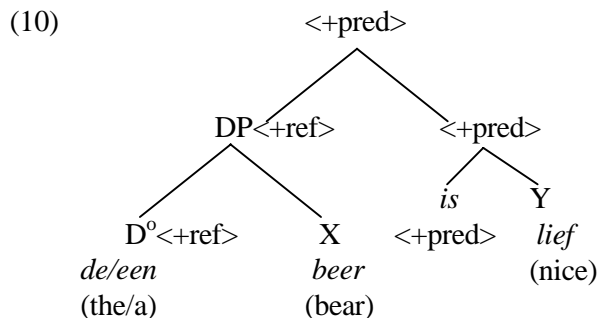
Dutch (Sarah, Van Kampen corpus)

a.	dit∅:	hondje weg	(1;10.5, w.96)	(this∅ (is) doggie gone)
b.	dit:	trein aan	(1;11.15, w.102)	(that (is) (comes) train on)
c.	dit∅:	huilen	(1;11.15, w.102)	(this∅ (to) cry)
d.	dit∅ niet:	bad zwemmen	(1;11.15, w.102)	(this∅ (is) not (in) bath (to) swim)
e.	dat niet:	lachen	(2;0.17, w.107)	(that (is) not (to) laugh)
f.	dit is:	zoeken	(2;0.17, w.107)	(this is (to) search)
g.	deze ook:	koud buiten	(2;2.18, w.116)	(this (is) also cold outside)
h.	hier:	is dat nou?	(2;3.16, w.116)	(here (is) (what) is that then?)
i.	dit i(s):	vliegtuig weg	(2;4.2, w.122)	(this i(s) airplane gone)

In the constructs in (2) with two content elements the content signs *beertje/mamma/deur/pop* ('bear/mummy/door/dolly') are used in the first position as the reference intended sign X. Admittedly, the content signs Y in (2) *slapen/dom/op* ('sleep/stupid/on') are atypical for child language as occupants of the first position, but they cannot be excluded in principle, see the attested cases in (3). Hence, the constructs in (9) do not necessitate a syntactic category N.



Categorization is delayed until the content categories $\langle +/ -N \rangle$ $\langle +/ -V \rangle$ are suggested by more specific syntagmatic properties. Suppose the child gets the problem of a three-word predication in (10), where articles *de/een* ('the/a') are used as prefixes to a referentially intended X.

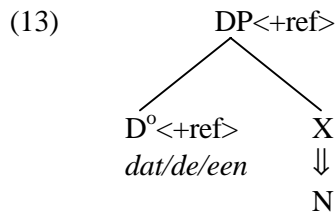


The pragmatic intention of using X as $\langle +ref \rangle$ in (2) is now expressed syntactically by means of D^0 , a non-characterizing referential functional category. This opens the possibility of a two-step bootstrapping procedure in (11) and (12).

(11) <+ref> ⇒ D (Bootstrapping:
from pragmatic intention to syntactic category)

(12) X<lex> ⇒ N/ D — (No bootstrapping: syntax intern)

The first step in (10) identifies the category D as a sign for <+ref>. The second step in (12) identifies a sense-bearing word X as having that reference. The significance of the lexical category N would be: “is bound to build referential structure with D”. Since, in adult Dutch, the demonstrative *dat* may appear in the subject position as an independent referential element D⁰ (without an N), *dat is een beer* ‘that is a bear’, it is to be recognized as having the referential index too.¹⁶ This will bring the child to (13).



If we look at the acquisition data of Sarah, we get the following picture for the reinterpretation of the deictic operator as D⁰. At first *dat* for presentationals (14)a broadens its use as simple utterance announcers as shown in (14)b, where a wh-question is announced by the deictic operator.¹⁷

(14) Dutch (Sarah)

- | | | | |
|----|-------------------|------------------------------|----------------|
| a. | dat: beer | (that: bear) | (from week 86) |
| | dat: zwemmen | (that: swim(ming)) | |
| | dat: leuk | (that: nice) | |
| b. | dit: is dat nou? | (this: (what) is that then?) | (week 116) |
| | deze: is dat nou? | (this: (what) is that then?) | |
| | hier: is dat nou? | (here: (what) is that then?) | |

Subsequently, three new phenomena, (15)a, (15)b and (15)c, in the acquisition of D⁰ marking seem to support each other in file 12 (week 120).

¹⁶ Category assignment is an idiomatic property of the individual lexical items. So here has to be an input controlled procedure for any category acquisition procedure. I won't start here all kind of possible scenarios to identify syntactic categories under all imaginable circumstances, since I am here busy with Dutch and English. To give an example, Russian, like many other languages, does not have articles, but it has demonstrative-like elements. The same demonstrative pronoun can be used independently, e.g. in presentationals, or attributively, added to an NP projection. Referential expressions in Russian need not have a demonstrative-like pronoun, it is an option, like in Dutch. However, referential expressions do need <+Case> in Russian and <+det> in Dutch. So, the acquisition procedure turns a referential intention into demonstratives or articles for Dutch, and the acquisition procedure turns referential intention into demonstratives and Case in Russian.

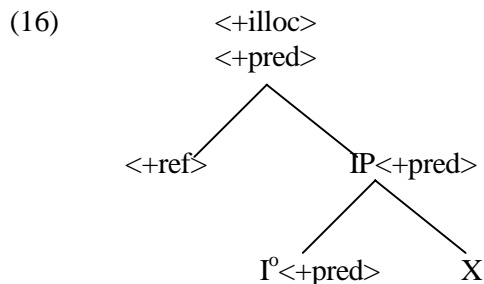
¹⁷ Next to demonstratives, child language uses deictic locative adverbs (*here/there*) in presentationals. Like English, the Dutch input doesn't allow deictic locative adverbs (*hier/daar*) to appear as D⁰. However, the use of deictic locatives as D⁰ is a possibility in other languages. Consider for instance the suffixes *-ci* and *-là* in French: *cette maison-ci* ('this house here') and *cette maison-là* ('that house there'), cf. Strawson (1974:97), Lyons (1977: chap 15, 1979).

- (15) Dutch (Sarah)
- | | | | |
|----|---|----------------------|------------|
| a. | “boe”, <i>zei de koe</i> | (“moo” said the cow) | (week 120) |
| | <i>dat is een wesp</i> | (that is a wasp) | |
| b. | <i>dat is een paddestoel</i> | (that is a mushroom) | (week 120) |
| c. | <i>ik wil [deze spelletje]_{DP}</i> | (I want this game) | (week 120) |

Firstly, the use of referential marking by D^0 has started to rise to >10% of all referential expressions at week 120, as will be shown in graphs (25)-(26) in section 5. See the examples [*een wesp*]_{DP} (‘a wasp’) and [*de koe*]_{DP} (‘the cow’) in (15)a. Secondly, the former presentational gesture word *dat* (‘that’) appears as a non-stressed pronominal subject, as shown by the example [*dat is een paddestoel*]_{IP} (‘that is a mushroom’) in (15)b. This must be the emergence of the EPP (Extended Projection Principle, i.e. obligatory subjects) for predications in Dutch. Thirdly, *dat* is for the first time used as a D^0 , see the example [*ik wil [deze spelletje]_{DP}*] (‘I want this game’) in (15)c. The common point in these three phenomena is the reinterpretation of the pragmatic sign for presentationals as a general sign for D^0 referential intentions, as has been sketched above. The reinterpretation does not exclude its use in presentationals.

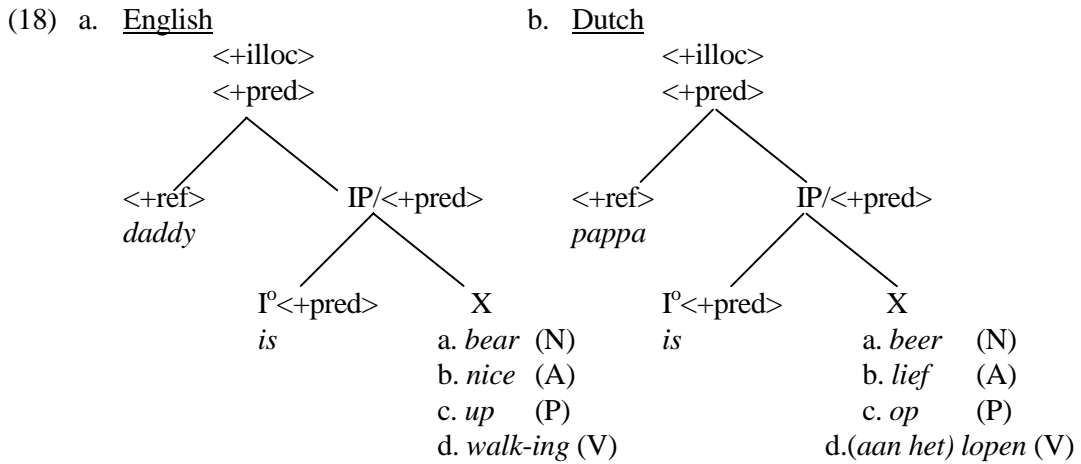
4 Bootstrapping V

The bootstrapping of $X = V$ may proceed in a similar way as the bootstrapping of N, using the context I^0 and some inflection morphemes, see (16) and (17). If a sense-bearing element X carries a predicational index, it should be accompanied by function-bearing constants I^0 .



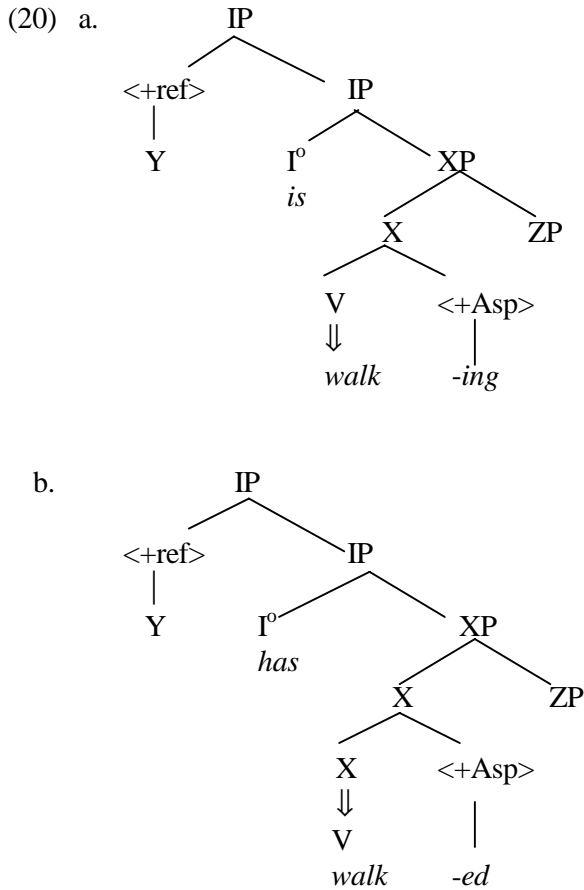
- (17) a. <+pred> \Rightarrow I (Bootstrapping: from pragmatic intention to syntax)
 b. X<lex> \Rightarrow V / I — (No bootstrapping: syntax intern)

If one looks at frame (17), though, the following objection may arise. The I^0 elements in the adult target language are the copula, the modal, the auxiliary and the finiteness on the verb. Characterization by a non-copula auxiliary or by <+fin> morphology may separate the V-class from other lexical classes. However, one of the members in I^0 is the copula *is*. The immediate application of (17) will now have the undesirable effect that all predicate heads are turned into V for Dutch as well as for English, see (18). The structure for Dutch in (18) abstracts away from directionality and verb movement (V-second).



The abstract structural context condition therefore needs to be narrowed down. For example, the English input distributions in (19) demonstrate that the tense auxiliaries I^o may induce a content category X \Rightarrow V if the context I^o includes the morphology on *-ing* / *-ed*, see 20.

- (19) a. X<lex> \Rightarrow V / [is — *-ing*]
 b. X<lex> \Rightarrow V / [has/is — *-ed*]



The acquisition picture for V-second Dutch is different, but the problem for rule scheme (17)b remains the same.¹⁸ The Dutch input distribution in (21) demonstrate that the tense auxiliaries I^0 may induce a content category $X \Rightarrow V$ if the context I^0 includes the morphology on (*is*) *aan het* ('is on the', present continuous) and *ge-* (past participle morpheme).

- (21) a. $X<lex> \Rightarrow V / [is \text{ aan het} \text{---}]$
 b. $X<lex> \Rightarrow V / [heeft/is \text{ ge-} \text{---}]$

An experiment carried out by Krikhaar & Wijnen (1995) showed that the Dutch present continuous structures like 'is *aan het* lopen' ('is on the V' = 'is walk-*ing*') was a very successful syntactic cue in Dutch child language for categorizing $X \Rightarrow V$. The triggering context abstract I^0 then takes the more specific form '*is aan het*' ('is on the'), cf. (18)d.

Clearly, the acquisition of the category V does not follow from any I^0 marking, but only from some (fairly frequent) I^0 types. These types are in addition accompanied by fixed inflectional elements (present/past participle; *te* ('to') +infinitive). Other predicate markers I^0 (i.e. the copula, modal verbs in Dutch) are less selective for the type of predicate that they can mark. See the types of modal constructions in (22) and (23) taken from the files of Sarah.

(22) Dutch (Sarah)

Illocution operator

a. *Modal operator + non-verbal predicate*

1) *Modal + N*

- (i)kwi(l) vogel (1;9.10, week 93) (wanna bird)
- (i)k moes (=moet) ∂ snottebel (2;1.10, week 110) (I-must ∂ snot)

2) *Modal + P*

- (i)kwi(l) uit (1;11.15, week 102) (wanna out)
- kan niet bij (2;4.2, week 122) ((I) cannot (reach) at)
- kan wel best uit (2;9.7, week 145) ((this) can indeed off)

3) *Modal + A*

- mag wel vies (2;4, diary) ((I) may indeed dirty)

b. *Modal operator + verbal predicate*

- hoef ∂ niet wassen (2;0.17, week 107) ((I) need ∂ not wash)
- mag ∂ buiten spelen (2;0.22, diary) ((I) may ∂ play outside)
- (i)kwi(l) leggen (=liggen) (2;1.10, week 110) (wanna lie down)
- (i)kmoe(t) slapen, hier (2;1.10, week 110) (I-must sleep, here)

¹⁸ The acquisition of the I-elements in Dutch is crucially related the acquisition of the V-second rule. It implies the awareness that any lexical $X<+V>$ can be turned into an independent I^0 due to a morphological element $<+fin>$, i.e. the awareness of an underlying pattern. The relation of $V_{<+fin>}$ with the (empty) predicate final position follows from a backtracking procedure for theta-assigners (Van Kampen 1997:chapt 3, p. 162). The backtracking procedure is confirmed by the fact that subordinate structures are acquired instantaneously, but cannot and do not appear before the backtracking of the V-second has been acquired. I will not go into that problem here, but see Evers & Van Kampen (2001) for a full and quantified analysis of the V-second acquisition scenario.

(23) Dutch (Sarah)

Independent predicate

- | | | |
|---------------------------|--------------------|-------------------------|
| a. ik wil/(i)kwi(l) | (2;2.18, week 116) | (I want/wanna) |
| b. ik hoef | (2;1.10, week 110) | (I need) |
| c. kan/hoef/(k)wi(l) niet | (various files) | ((I) can/need/want not) |
| d. mag/(i)kwi(l) wel | (various files) | ((I) may/want indeed) |

In the same vein, *wanna* constructions in English child language mark the modal status of verbal as well as non-verbal predicates (Gruber 1975, Powers 1996a).¹⁹

The partial selective function of the I^o elements, sometimes obligatory followed by an X<+V>, sometimes not, has a parallel with the D^o elements, as shown in section 3. The articles *de/een* ('the/a') invariably select an X<+N>, cf. (10). The demonstrative *dat* ('that') can appear as an independent element (as a presentational operator or as an independent referential subject), but will also appear later on as attributive demonstrative in the position of the more neutral article, cf. (13).

5 Graphs for the acquisition of D^o marking and I^o marking

5.1 Provisional graphs

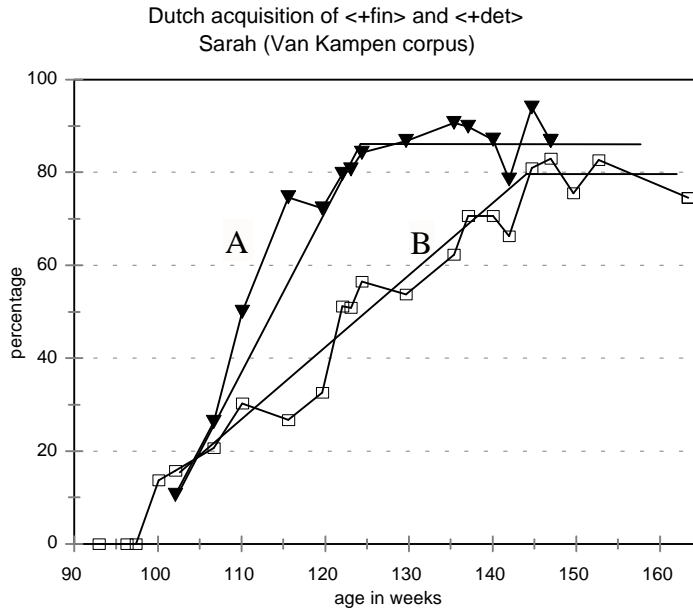
The maternal input marks nouns with referential function by a D^o. Sarah's mother lives up to the D^o norm in obligatory contexts for >95%.²⁰ In the same vein, subject-predicate constructions in Dutch root sentences are marked by I^o elements in V-second position. This holds for >90% in the adult input. The graphs in (24) show how two-year old Sarah acquired the skill for the obligatory I^o marking within 20 weeks and the obligatory D^o marking 25 weeks later. I will consider a grammatical feature acquired when the child is within 10% of the adult norm. This point in time will be referred to as acquisition point. The point in time where the child crosses the 10% line irreversibly is referred to as eureka point. The acquisition graph between eureka point and acquisition point is modeled in (24) by a linear function. For this simplification, see Evers & Van Kampen (2001).²¹ As a matter of fact, the D^o graph (B) seems a bit disappointing. It starts to settle around the 80% line, whereas the I^o graph (A) starts to settle at a persistently higher level. See the respective high horizontals of A and B in (24).

¹⁹ Gruber (1975) used the term 'performatives' for such predicate operators, as they were accompanied by a pointing or reaching gesture. Within the present view, one may reformulate this as follows. The deontic operators in early child language are used in deictic contexts to express subjective modality exclusively (see for subjective modality Lyons 1977: chap. 17). The first person pronominal subject *ik/I* is often absent or cliticized to the modal operator. In Van Kampen (1997:106) the first person subject drop was characterized as situation bound 'confession mode', a mode comparable to the imperative.

²⁰ The use of a D^o is obligatory with singular count nouns and with definite plural nouns. The D^o system in Dutch makes use of zero signs for mass nouns, for proper nouns, for indefinite plurals and for idioms.

²¹ The graphs for Dutch Sarah are constructed from the Van Kampen corpus in the CHILDES archive, from files 04-30, recording Sarah's first acquisition of Dutch between 1;9.10 (week 93) and 2;10.18 (week 163). One-word utterances, partially intelligible, incomplete or unclear utterances, imitations, immediate repetitions and formulaic routines were left out.

(24)



Number of examples graph A: 85 – 114 – 83 – 150 – 101 – 178 – 122 – 210 – 168 – 220 – 114 – 453 – 201 – 223 – 215

Number of examples graph B: 11 – 9 – 23 – 33 – 63 – 63 – 34 – 53 – 43 – 86 – 53 – 85 – 215 – 158 – 58 – 82 – 89 – 121 – 53 – 43 – 52 – 71

Graph A: Sentences that realize finite verbs in first or second position in \geq two-word utterances

Graph B: Noun phrases that realize determiners in obligatory contexts in \geq two-word utterances

The I^0 marking of the subject-predicate constructions starts at the same moment as the D^0 marking of nouns. This is interesting if ‘reference’ versus ‘predication’ is the major opposition of category-neutral sign application by the young speaker. Nevertheless, the subsequent developments for D^0 and I^0 are not parallel. The acquisition of the V-second I^0 properties takes less than half the time (20 weeks) needed for the reference D^0 properties (45 weeks). The more robust and merely formal nature of the I^0 marking in Dutch as compared to the D^0 marking may underlie this difference. Consider the following relevant factors for the delay. In the first place, there is but one I^0 marking for each sentence/predication (V-second), whereas a sentence/predication may contain several signs that are to be marked for reference. Most important probably is a second factor. The finite I^0 is exclusively <+present> during the acquisition period. I^0 marking can be set without reliance on tense oppositions. Moreover, Sarah’s mother hardly used past tense sentences. By contrast, the D^0 switches between <+definite> and <-definite>. For that reason, the D^0 marking requires more awareness of discourse and situation conditions. And although two-year old Sarah can make such decisions at week 120 (cf. the examples in (15)a) -, they will take more of her cognitive energy. The tenability of this explanation will be shown in the next section. The same explanation for the later appearance of the D^0 in Romanian child language is given in Coene & Avram (2001). Coene & Avram weaken that argument somewhat. They argue that child language at this stage does make a distinction between present and past tense. They point to the early appearance of past participles as realizations of past tense. To my mind, the past participle has little or no credentials for I^0 status. It can be characterized as a <+state> predicate, opposing a <+event> infinitive (*Laura (ge)daan* ‘L. done’ versus *pappa doen* ‘daddy do’). The idea that the past participle/perfect develops from aspectual towards temporal use is present in several studies (see for instance Bronckart & Sinclair 1973 for French, Bloom, Lifter & Hafitz 1980 for English, and

Gagarina (2000) for Russian child language).²² Both types of predicate will require I° <+fin> marking. The acquisition of that element <+fin> will not be completed by a <+/-past> opposition in case of event or state predicates.

Hence, the main suggestion here is that the Dutch D° is more difficult to acquire than the Dutch I° because the oppositions for the D° system are involved in discourse reference tracking, whereas the I° system is not yet burdened with discourse marking functions.

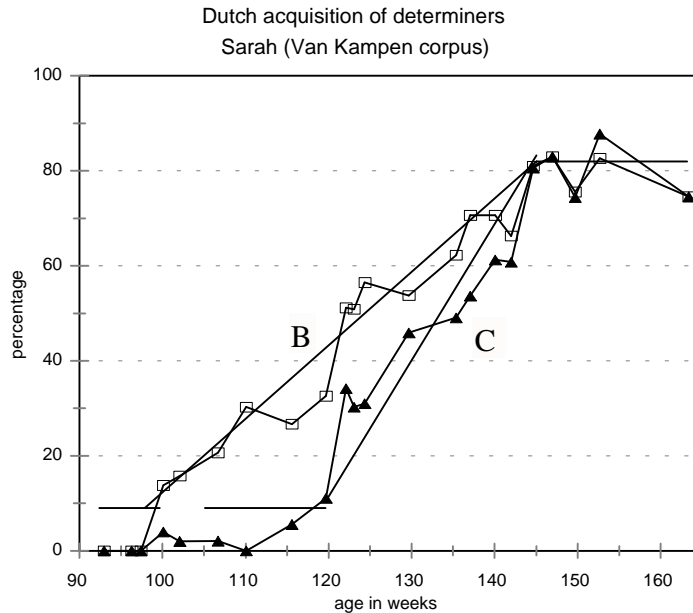
5.2 Revised graphs

The graphs in (23) seem to indicate that the use of determiners grows considerably slower than the use of V-second. There is, though, an interesting difficulty with the measurement of D° acquisition in Dutch. This difficulty does not arise with the measurement of V-second. Its full recognition will lead to a significant reinterpretation of the delay for D° of 25 weeks, cf. (24). The definite articles *de* ('the') and *het* ('the' before neuter N) and the indefinite article *een* ('a') are subject to rules of phonetic assimilation in (adult) Dutch. The consequence is that they are often only slightly different from a schwa (ə). At the same time various studies have reported that the language acquisition procedure starts the introduction of a functional category by marking the position with a schwa (ə), a kind of placeholder. One might say that any functional category is at first introduced as a tentative <F?> with categorical and semantic constraints (Evers & Van Kampen 2001). Only after the structural constraints on <F> have become clear, the learner proceeds by filling in the appropriate phonological form. Graph B in (24) now was set up by contrasting for each file the percentage of bare N versus the percentage of N preceded by a full or schwa determiner. Since in the babbling of Sarah one could hardly distinguish a schwa from a fully realized article, especially during the first half-year period, the acquisition graph for the Dutch D° might also be set up in a more cautious way. One might contrast for each file the percentage of bare N versus the percentage of N preceded by clearly realized articles and leave all forms ə+N out as non-decidable. This results in graph C in (25) below. The systematic use of definite determiners does not arise before week 122. Graph C then measures the growth of the opposition <+definite> versus <-definite> D°. The beginning of the rise of graph C at week 120 coincides with the point Sarah starts making a distinction between definite and indefinites (cf. (15)a). The earlier rise of graph B must then be due to 'dummy' D°.²³ The comparison of graph B and C in (25) shows that graph C begins to coincide with graph B at the acquisition point (145 weeks). This means that the underspecified, provisional schwa-determiner becomes marginal (<10%) at the acquisition point of a D° for all referential expressions, as expected by the present analysis.

²² It has been claimed (e.g. De Houwer 1997) that English children acquire the simple past tense before the past participle/perfect tense, whereas the reverse would hold for Dutch/German children. I do not see that. English morphology makes in principle no distinction between past tense and past participle, so neither will the child. It is more likely that all children start with a past participle as a state-assigning predicate. The tense-oppositions will not enter the system before the acquisition of <+fin> and <+tense> about half a year later on. See Evers & Van Kampen (2001).

²³ The acquisition of the determiner in English follows the same route. At first, the English child realizes 'dummy' D° only. S/he does not use of definite determiners and *a* is not realized as *an* yet (the child says *a eye*, instead of *an eye*).

(25)

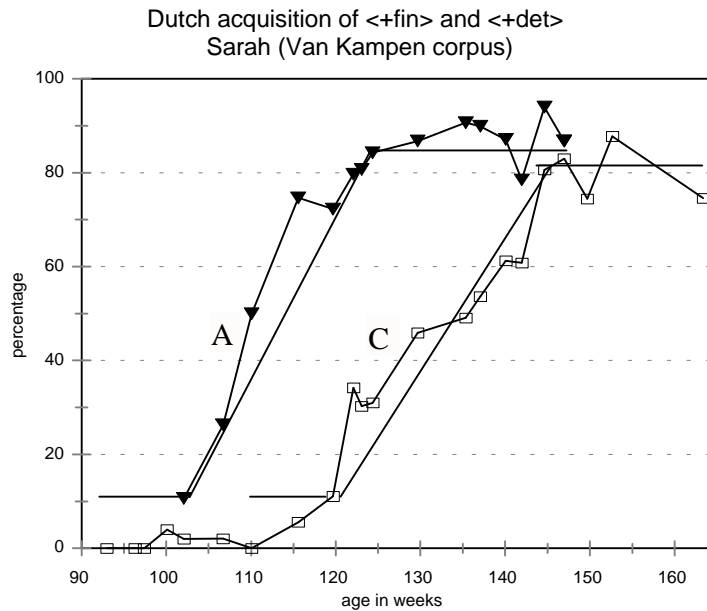


Graph B: Noun phrases that realize <+det> in obligatory contexts in \geq two-word utterances: ∂ determiners included

Graph C: Noun phrases that realize <+det> in obligatory contexts in \geq two-word utterances: ∂ determiners not included

I propose that graph C is the best representation of D^0 (determiner) acquisition for Sarah. If so, the D^0 delay of 25 weeks is due to a delay in the eureka point. The measurement problem did not arise for graph A. The schwa placeholder for the later V-second marking does appear in a certain number of cases, but there is no phonetic assimilation rule such that one might hesitate whether the placeholder is maybe the reduced form of a real finite verb. So the small group of schwas as V-second predicates has been counted as predicates not marked by a finite verb. Therefore the acquisition period of V-second ranges from week 100 to week 120. The systematic D^0 marking and the systematic I^0 marking now both appear during an acquisition period of roughly 20 weeks, see (26).

(26) Acquisition graphs of D^0 and I^0 revised



Graph A: Sentences that realize finite verbs in first or second position in \geq two-word utterances

Graph C: Noun phrases that realize determiners in obligatory contexts in \geq two-word utterances

The lower high horizontal of C as compared to A in (26) remains unexplained. Ideally, one would like to claim that a persistent 15% failure to reach the adult norm must be due to specific structural positions, e.g. subjects do well, but objects and PP adjuncts continue to lag behind (cf. Baauw, De Roo & Avrutin 2001). An alternative story would be to relate the 15% failure to the definiteness effects of a topic/non-topic distinction (Schönenberger et al. 1997, Van Kampen & Wijnen 2000). There are fluctuations but none of them impresses me as a sufficient support for such or any other explanation at the moment.

5.3 Number specification

Hoekstra & Hyams (1998) propose a common source for the acquisition of D^0 as well as I^0 . Number specification would be the underlying core distinction. The graphs in (26) show that this is not supported by the facts (see also Baauw, De Roo & Avrutin 2001).

Hoekstra & Hyams (1998) counted the occurrences of bare noun subjects versus the occurrences of specified DPs in both utterances with a <-fin> V and utterances with a <+fin> V for two Dutch children aged 2;3-3;1 (Hein) and 2;7-3;2.13 (Niek). See the reproduction of their table 13 in (27).

(27) Distribution of overt subjects for Dutch children (Niek and Hein)
 (from: Hoekstra & Hyams 1998: table 13)

a	Number of occurrences				percentages		
	b	c	d	e	f	g	h
		<-fin> V	<+fin> V	total	<-fin> V	<+fin> V	total
specified DP	overt D ^o – N	4	382	386	1% (4/386)	99%	100%
	plural marking on N	2	180	182	1% (2/182)	(382/386)	100%
	pronoun	169	4.407	4.57 6	4% (169/4576)	99% (180/182) 96% (4407/4576)	100%
unspecif. DP	bare N	28	423	451	6% (28/451)	94% (423/451)	100%

The differences in percentages in (27)f versus (27)g are so striking as to become suspicious. A somewhat closer look reveals that the differences are due to biased choices in sample as well as biased choices in selection period. Adding up the numbers in column (27)c yields 203 <-fin> constructions. Adding up the numbers in column (27)d yields 5.392 <+fin> constructions. As far as the biased choice of sample is concerned, the sample is overwhelmingly <+fin>. Hence, it makes no sense at all to ask with Hyams & Hoekstra what support D^o categories may have given to the rise of the <-fin>/<+fin> distinction, since the sample is overwhelmingly <+fin> anyway. The answer can have no relevance to D-marking relations whatsoever. Even if a precocious child would have all his D-marking according to the adult norm from the beginning on, and had come up with, say, 30 cases of overt D^o, instead of the present 4, reducing his present unspecified DP from 28 to 2 (see column (27)c), he would still get bad marks from Hoekstra & Hyams: some lousy 7% (30/412 (=382+30)) as compared to the 93% (382/412) of the <+fin> constructions.

It may seem to make more sense to ask for the distribution of the four D-constructions in early <-fin> and <+fin> sentences, see the recalculation in (28). The picture at least changes completely.

(28) Recalculation of the table in (27)

a	Number of occurrences			percentages	
	b	c	d	e	f
		<-fin> V	<+fin> V	<-fin> V	<+fin> V
specified DP	overt D ^o - N	4	382	2% (4/203)	7% (382/5392)
	plural marking on N	2	180	1% (2/203)	3% (180/5392)
	pronoun	169	4.407	83% (169/203)	82% (4407/5392)
unspecif. DP	bare N	28	423	14% (28/203)	8% (423/5392)
	total	203	5.392	100%	100%

As one may see in column (28)e, the unspecified DP in <-fin>V sentences holds some ground (14%). This keeps the overt D^o (2%) and the plural marking (1%) lower than the corresponding percentages (7% and 3%) in the <+fin> sentences of (28)f. Yet, that is not relevant either. The period is still biased. If one would extend the period for selection of construction, say from the second to the twenty-first birthday, the percentages for grammatical N-marking will become indiscernible from the percentages in the adult language for <-fin> as well as for <+fin> constructions. This only proves there is a period of full competence. If, by contrast, the period is shortened and only alternations before two and a half year are selected, than undoubtedly the percentage of unmarked N will grow and D-marking will be lower. This proves that there is an acquisition period of N-marking, but little more. If one really wants to be informed about the relation between D-marking and I-marking, there is no alternative but to construct longitudinal graphs for individual children.

An expectation along the lines of Hyams & Hoekstra, might have been that D-insertion with nouns is acquired along with Spec-V<+agr> constructions. Baauw, De Roo & Avrutin (2001) looked at this correlation in the development of two Dutch children. They divided the acquisition of <+fin> marking in three periods: 1st period of <30% <+fin>; 2nd period of 50% <+fin>; and 3rd period of >70% <+fin>. As predicted by my graph in (26), there was no clear correlation between I^o<+fin> and D^o<+det>. The two children showed a preference for D^o<+det> in I^o<+fin> utterances in the last period (>70% <+fin> marking), but only for determiners with nouns in general and not for determiners with subject nouns. My own counts, below in (29), do not reveal such a preference. The effect is not particularly striking either way, and I think the measurement variations are due to random noise.

(29) Percentages of <+det> before nouns in finite and non-finite utterances (∂-det not counted)

(Sarah, Van Kampen corpus)

age in weeks	<-Vfin>		<+Vfin>	
100-102	4/66	6%	0/8	0%
107-110	9/69	13%	0/12	0%
116-120	5/34	15%	6/38	16%
122-123	6/35	17%	21/55	38%
125-129	21/40	53%	62/148	42%
132-136	11/19	58%	52/103	50%
137-140	16/21	76%	44/78	56%
142-145	17/24	71%	110/149	74%
147-150	18/21	86%	44/57	77%
153-159	15/16	94%	81/101	80%

In short, the quantitative argument by Hoekstra & Hyams is invalidated by elementary mistakes in the selection of data. Their research interest as such was nevertheless a very interesting one. As I see it, they wondered whether licensing by Spec-head agreement plays a demonstrable part in the acquisition of I^o<+fin> and D^o<+det>/<+pro>. To put it somewhat informally, are the PF reflections of argument licensing in child language an adornment added later on or is it the very core issue of grammar? My guess would be ‘an adornment added later on’. I tend to see I^o and D^o as autonomous deictic markers of

predication and reference that come in to construct discourse cohesion rather than sentence cohesion. They may pick up their later syntactic function as well, after some fine-tuning. So, real morphological agreement and Case-marking will be acquired later on. The point of grammar is not sentence cohesion, but discourse cohesion. The next section will show how the acquisition of D°/I° marking enables the child to build up discourse cohesion.

6 From situation-oriented towards discourse-oriented language

The acquisition of D° and I° in early child language, as pictured in the graphs of (26), imposes a systematic use of the parts of speech distinctions N/V. The mechanisms for this effect have been given in the rule schemes in (11)-(12) and (16)-(17). The distinction of N/V becomes systematic due to the DP and IP frame they fit into. The N/V structured lexicon will probably support the recognition and acquisition of all kind of lexical semantic contrasts. See also the comparable suggestions by Maratsos (1982), Maratsos & Chalkley (1983), Gleitman (1990) and O’Grady (1997/1999) and the acquisition simulation by Brent (1994).

The D/I-marking has a second major effect and this effect may also enhance the acquisition of the lexicon. The graphs in (26) reach the near native adult level, the so-called acquisition point (Van Kampen 1997), in respectively week 120 and 145. From that point on, the various content distinctions made in a conversation are marked deictically by the D°/I° devices. Also from that point on, and not earlier, it becomes possible to refer back to these distinctions in sentence discourse. The grammatical oppositions by means of which one may refer back to previous discourse make up discourse grammar.²⁴ The underlined parts in (30) show the grammatical underpinnings for discourse cohesion.

- (30) Snow-white_i was offered a_j poisoned apple
 (and then) she_i took a hasty bite from the_j apple (it_j, *an_j)
 (and then) it_j (*an_j, the_j) got stuck in her_i throat
 (and then) she_i fell on the floor (*a, *it)
as if she_i were dead

The passage relates the individuals *Snow-white* and *an apple* to a sequence of states and events. Consequently, the relevance of Snow-white and her apple changes as the story evolves. It is possible, and usually preferable, to refer back to Snow-white and the apple by means of discourse anaphora (*she, it, her*) and definite determiners (*the, *an*). The sequence of events is indicated in discourse by tense elements and further qualified by modals and various connectives (complementizers and adverbs for cause, consequence, condition, circumstance, and so on). The tense in the passage above not only indicates “fiction, no relevance to the actual situation”, but also “temporal sequence”. If the story switches towards past continuous, the competent listener should derive simultaneity of events, cf. (31).

²⁴ Discourse is intended here in the narrowest sense of the word: using grammatical devices to refer back to earlier constructions in other sentences. It is not meant in the broad sense of “the cognitive orientation presupposed in a conversation”. The cognitive orientation of discourse in its broadest sense is a cultural educational result brought about by discourse in its narrowest sense.

- (31) Thereafter, Snow-white was being placed under a glass coffin window and the prince and his horse were wondering whether a true princess was not bound to stay under the glass jar forever.

Heim (1982) proposed to represent discourse cohesion that follows from D^0 marking by means of file cards. The D^0 file cards preserve an identity through the events and successive states that characterize time. Avrutin (1999) has proposed to introduce file cards as well for events. These event file cards will serve to represent the contribution of I^0 marking to discourse cohesion. Since the I^0 marked distinctions simply are qualified parts of time/modality, they are, I suppose, outside updating as time goes on, whereas the D^0 file cards do need just that. I assume that the I^0 file cards are needed to represent the event sequencing and its modal qualifications as discourse is built up.

File cards model cohesion as it arises from the D^0 markings and I^0 markings of a language. Therefore, I take it that a communication system without these markings will lack marked discourse cohesion and make no use of file cards whatsoever. Systematic D/I-marking (<+/-def>; <+/-past>; <+/-asp>) is not needed for a language that makes no reference to discourse, hence needs no file cards. The pre-occupation of most or all language types with systematic D/I-marking shows how these systems are set up for discourse referencing, and need file cards to keep track of discourse cohesion.²⁵

Children start with deictic reference only as a kind of situational gesturing (cf. Lyons 1977: 648, 1979). If there is situational ‘salience’, deictic oppositions are needed only by exception, such as when a crucial distinction is to be made between *this* and *that*, or between *now* and *then*. This exceptional case does not occur in the early language of the child (Clark and Sengul 1978, Fletcher 1985). Demonstratives *dit/dat* (‘this/that’), place adverbials *hier/daar* (‘here/there’) and time adverbials *nou/dan* (‘now/then’) appear at first as illocution operators. The acquisition of I^0/D^0 introduces later a deictic <+/- proximate> opposition that leads to discourse oppositions of deictic reference (*this/that; here/there*) and deictic tense (<+/-past tense>; *now/then*).²⁶ Before the child applies the D^0 and I^0 oppositions in a discourse appropriate manner he/she must lack an effective grip on discourse cohesion. The discourse use of D/I-marking is highly language-specific and hardly learnable after puberty, e.g. the aspectual system in Russian. For that reason, the acquisition of the D^0/I^0 implies that there is such a thing as the acquisition of discourse grammar, i.e. the acquisition of a file card system. And it implies in fact that the acquisition of the D^0/I^0 and the acquisition of the discourse grammar are one and the same thing. It stands to reason that the perception of discourse grammar enlarges the opportunity to build up lexical distinctions. By contrast, as long as the child does not yet apply the D^0 and I^0 devices, its construction of a coherent

²⁵ Languages may differ in having Case and/or articles, Aux and/or inflection, and various distributional restrictions. So one can imagine all kind of possible scenarios a language can keep track of discourse cohesion.

²⁶ From a more speculative point of view, I propose that the very possibility of entering a discourse orientation is dependant on a grammatical system that infuses its phrases with deictic markers for later reference. This point of view is also advanced in Lyons (1979). It explains the fact that natural language abounds in deictic marking that (to the beginning speaker at least) seems situation-redundant. Grammatical systems seem inherently be designed to break the restrictions of situational saliency at any moment. This tallies well with the major point of this paper: children do not attend to discourse relevance, until they have mastered the grammatical I^0/D^0 devices. They do not venture earlier into discourse, because without deictic D^0/I^0 marking one cannot handle information that way.

discourse must be poor and he/she will derive little acquisition support from discourse. The growing reliance of the child on discourse grammar is exemplified by opposing the following two dialogues, taken from Crystal (1987:246), quoting Fletcher (1985:64, 92).

(32) English

a. Sophie 2;4.28

child: ball. kick. kick. daddy kick

mother: that's right, you have to kick it, don't you.

child: mmm. um. um. kick hard. only kick hard. our play that. on floor. our play that on floor. now. our play that. on floor. our play that on floor. no that. now.

mother: all right.

child: mummy come on floor me.

mother: you tip those out.

child: mmm. all right.

mother: that one broke. when did that happen?

child: Muffy step on that.

b. Sophie 3;0.4

child: Hester be fast asleep, mummy

mother: she was tired.

child: and why did her have two sweets, mummy?

mother: because you each had two, that's why. she had the same as you. oh dear, now what?

child: daddy didn't give me two in the end.

mother: yes, he did.

child: he didn't.

mother: he did.

child: look, he gave <one to> [/] two to Hester, and two to us.

mother: yes, that's right.

child: why did he give?

mother: (be)cause there were six sweets. that's two each.

The difference in discourse cohesion between (32)a and (32)b is a strong, but merely intuitive impression. I propose that a quantitative measure of discourse reliance can be found for D⁰ marking by determining the percentage of referential expressions that refer back into discourse (announcing familiar information, previously mentioned or clear from the situation) by means of definite determiners or discourse anaphora (*he/she/it/him/her* etc.). Reliance on discourse anaphora will of course vary with the type of conversation in the file. Assuming a balanced even-handed conversation between mother and child, it is plausible that the percentage of discourse references reached by the child will at first trail the percentage used by the mother, while eventually the child will catch up. The same will hold for grammatical devices that construct event sequencing. The amount of formal discourse structuring eventually rises to the level of the adult conversation partner.

There are many things to be figured out, but the rise of discourse grammar can be measured. At first the Dutch child uses the following referential markings: a) a schwa determiner, but no <+def>/<-def> opposition; b) 1st and 2nd person pronouns; c)

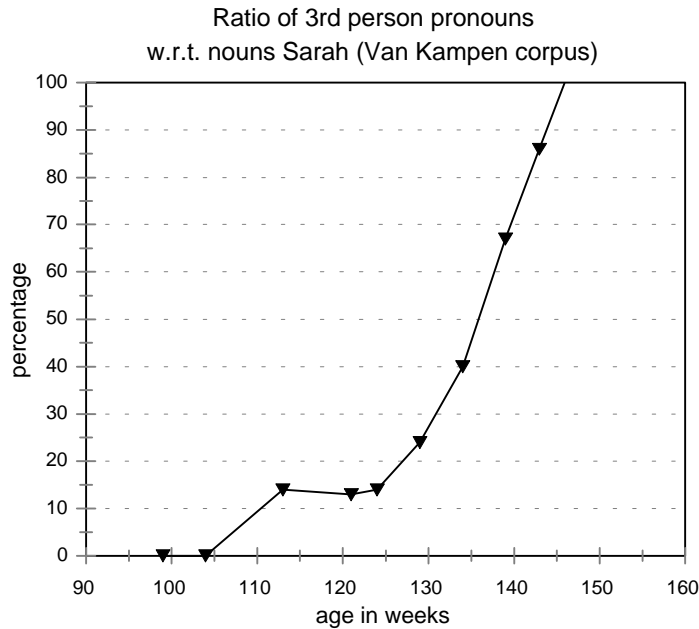
demonstrative pronouns (*dit/dat* ‘this/that’), but no 3rd person pronouns (*hij/ie/zij/ze* ‘he/she’). These three markings fit into the picture of deictic reference as strictly situational and non-discourse bound. As shown in section 5.2, the D⁰ positions appear at first as mere place-holders represented by a schwa that does not distinguish between <+/-definite> (see also Van Kampen & Wijnen 2000). Sarah starts using definite determiners marginally at week 116, but the systematic use of definite determiners does not arise before week 122.

The same development from non-discourse relevance towards discourse grammar relevance holds for the use of pronouns. The early use of 1st and 2nd person pronouns is easily explained. They are completely situation-oriented deictic pronouns. The 3rd person pronouns in Dutch are interesting. Dutch is a topic-oriented language. Strictly situation-relevant 3rd person pronouns are preferably introduced by the demonstrative pronoun (*die/dat* ‘this/that’). The use of the personal pronouns for 3rd person (*hij/ie/ze/zij/hem/haar/hen* etc, ‘he/she/him/her/they/them’) is mainly discourse-oriented (previously introduced 3rd person <+/-male>, <+/-plural>). It is true that the opposition between the uses of the two pronouns is a matter of preference not a prescript, but by constructing a graph for the rise of the discourse oriented *hij/ie* as a percentage of the amount of referential expressions, one can measure the growing ability of the child in discourse reference tracking. In the first period, Sarah uses only demonstratives. Moreover, she uses the demonstrative exclusively as a presentational marker, without saliency, i.e. without making a distinction between proximal (*deze/dit* ‘this’) and distant nouns (*dat/die* ‘that’). Only later she starts using 3rd person pronouns. The graph in (33) shows how Sarah over the weeks catches up with the 3rd person discourse anaphors of her mother.^{27 28}

²⁷ Each measuring point in the graph in (33) represents two consecutive files from the Sarah corpus. For each measuring point, the ratio of the pronoun-density in the speech of the child is measured as a percentage of the ratio of the pronoun-density in the speech of the mother (within the same files).

²⁸ In Dutch child language 1st and 2nd person pronouns (*ik/mij/jij doen* ‘I/me/you do(ing)’) and demonstratives (*die doen* ‘that (one) do(ing)’) appear as subject of a non-finite sentence, but never a third person pronoun (*hij doen* ‘he do(ing)'). At least I did not find any examples.

(33)



At week 145 Sarah has acquired the $\langle +\text{def} \rangle / \langle -\text{def} \rangle$ D^0 marking of nouns (>80%), see graph C in (26). Curiously enough, it is the same week that her use of discourse anaphors *hij/ie/hem/ze/zij/het* reaches the level of her adult conversation partner.²⁹ Clearly, we have here an interesting measure for the growing ability of the child in handling grammar. There are a variety of discourse relevant devices that develop as soon as the D^0 marking is acquired. It is for instance worthwhile to notice that the mistaken addition of the definite article to proper names (e.g. *de pappa* ‘the daddy’) appears momentarily at this point (week 145) as well.

There are also a variety of discourse relevant devices that develop as soon as the I^0 marking is acquired, e.g., the use of adverbs as discourse relevant tense indicators, the use of verbal tense and aspect oppositions and the development of epistemic modals from root modals. Time adverbials like *nou/nu* (‘now’) and *dan* (‘then’) seem to appear earlier, but at first they function only as markers of turn takings in the conversation, cf. section 2 and Evers-Vermeul (2000). The element *nou* solicits a reaction from the conversation partner and the element *dan* merely claims the right on a taking a turn. Only later these elements are used for deictic/discourse marking, when the child makes a distinction between *nou/nu* (‘now’) and *dan/toen* (‘then’).³⁰

²⁹ See also Hamann and Plunkett (1998). They notice that Danish child language also shies away from third person pronouns for a well-marked period. Hamann and Plunkett suggest as well that children might initially lack the necessary discourse anchoring. The option to use a demonstrative instead of a 3rd person pronoun does not exist in non-topic languages like English or French (Van Kampen 1997: 112). Nevertheless, a comparable lack of discourse anchoring can be shown for English and French pronoun acquisition. E.g. Karmiloff-Smith (1981) used elicitation experiments to show that French children initially use 3rd person pronouns in situation bound contexts only.

³⁰ In a study on the acquisition of connectivity Berman (1998) argues that *ook/auch* in German/Dutch function as precursors of connectivity. She rightly observes that every language may have his typical immature precursors of discourse markers, e.g. *more/again* for English child language. Berman (1998) claims that children connect clauses by juxta-positioning the adverbials with an additive function, like

If we look at the acquisition data of Sarah, we get the following global picture. The first attested cases of adverbials used by Sarah as connectives, is in week 122, at which time she also starts using verbal past tense, see the examples in (34). See for the acquisition of connectives in Dutch Evers-Vermeul (2000), in English Bloom et. al. (1980).

- (34) (Sarah, week 122)
- a. was ik (ge)vallen (was I fallen)
 - b. toen ∂ ik huilen (then ∂ I cry)

At this point, week 122, Sarah has already completely acquired the V-second I° properties (>80%), see graph A in (26).

7 Conclusions

The present perspective departed from bare projections of grammatical features that indicate the pragmatic intentions of the sign (De Saussure 1915, Lyons 1977/1979). The acquisition of grammar (I°/D°) subsequently results in cognitive systems like {parts of speech (N/V), file card structure}. *So first there is grammar and then a cognitive system.* This perspective differs from ‘cognitive’ views as expressed by Pinker (1984: semantic bootstrapping) and Avrutin (1999: file cards by presupposition).

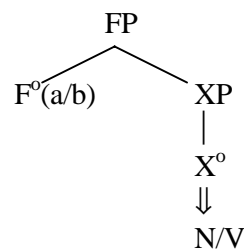
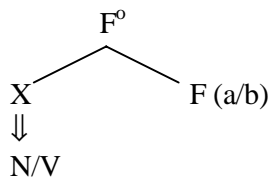
The present paper sketched a strategy for bootstrapping the lexical categories N and V from functional categories for the pragmatic intentions <+ref>, <+pred>. The differentiation between N and V cannot take place before <+ref> and <+pred> are expressed by functional categories D° and I° respectively. The functional categories are highly frequent and for that reason good bootstraps for the set of lexical content items. See the bootstrapping procedure in (35).

- (35) *Bootstrapping procedure*
- a. <+ref> ⇒ D / [— X]
 - b. <+pred> ⇒ I / [— X]

morphological context

or

phrasal context



It has been argued that the pre-existence of D° and I° elements themselves went back to illocution operators for presentationals and wish/command.

ook/again. The claim made here extends this. There could be a development from a-syntactic illocution marker (situation bound), towards sentence adverb (implicit discourse connection), towards syntactic connective (discourse bound).

The advantage of the procedure sketched above as compared to a cognitive semantic procedure that derives content categories from general notions, like ‘thing’ (N) or ‘action’ (V) (Pinker 1984), must be that functional categories are structurally identifiable without reliance on the ‘thing’/‘notion’ difference between the sense-bearing items, and they are reaffirmed all the time. These are the very properties one would like for ‘bootstraps’. Content words (lexical items), by contrast, have a lower text-frequency, a less outspoken distribution and their semantic properties are notoriously elusive.³¹ According to recent proposals, all grammar resides in functional morphemes (Borer 1984 and subsequent work, Marantz 1997, Chomsky 1995, 1999). Hence, the proposal above has the good fortune to fit in with recent theoretical views about grammatical categories.³²

The syntactic signs D^0 and I^0 identify distinctions that the speaker may refer back to by the grammatical devices of discourse grammar. Hence, the acquisition of the D^0/I^0 system is not only the key to the parts of speech system (N/V), but also the introduction of discourse grammar. Following this perspective, I have proposed that the acquisition of the D/I-system is to be considered as the introduction of discourse grammar as modeled by file cards for things and events (Heim 1982, Avrutin 1999). This view on D/I-marking explains why phenomena of discourse grammar appear in early child language right after the acquisition of systematic D/I-marking. Further, it explains why the various language types insist on the systematic, almost invariably, application of D^0 and I^0 marking. The system is designed for discourse grammar.

Acknowledgments

The research reported in this article was supported by a grant (355-70-009) from the Netherlands Organization of Scientific Research (NWO) and carried out at the Utrecht institute of Linguistics OTS. A version of this article has been presented in Van Kampen (1997, 2000a, 2001b). I would like to thank the members of the UiL OTS acquisition group for some skeptical notes, especially Sergey Avrutin, Sergio Baauw and Aafke Hulk. I received systematic feedback from Arnold Evers. The empirical results reported here are largely due to Rianne Vlooswijk, who collected the data for the acquisition of determiners and helped me to categorize them.

³¹ The same line of reasoning was perceived earlier by Buszkowski (1987). Buszkowski defined an algorithm for finding a categorial grammar from data that consist of functor-argument structures (Buszkowski 1987; Buszkowski & Penn 1990). Functor-argument structures are like phrase-structures without lexical category labels. The Buszkowski algorithm assigns variables to arguments (roughly: lexical categories). Subsequently, the arguments are labeled on the base of their configuration with a functor. The mere presence of a formal modal of learning does not of course imply that it will be applicable to the real life procedures of first language acquisition, but it is encouraging.

³² O’Grady (1997, 1999) is somewhat similar to Van Kampen (1997, 2000a, present paper) in its attention to minimal structures as a cause for category formation. His analysis diverges in two ways from the present one. Firstly, it is based on a simple categorial grammar without functional projections, i.e. it does not take (35) as a starting point. Secondly, there is no reference to actual acquisition data.

References

- Avrutin, S. (1999) *Development of the Syntax-Discourse Interface*, Dordrecht/Boston/London: Kluwer Academic Publishers.
- Baauw, S., E. de Roo and S. Avrutin (2001) 'Determiner omission in language acquisition and language impairment: syntactic and discourse factors', paper presented at the Boston Conference on language development, Boston University, November 2.
- Berman, R.A. (1998) 'Typological perspectives on connectivity', in: N. Dittmar and Z. Penner (eds.) *Issues in the Theory of Language Acquisition*, Bern: Peter Lang, 203-224.
- Bloom, L., M. Lahey, L. Hood, K. Lifter and K. Fiess (1980) 'Complex sentences: acquisition of syntactic connectives and the semantic relations they encode', in: *Journal of Child Language* 7, 235-261.
- Bloom, L., K. Lifter and J. Hafitz (1980) 'The semantics of verbs and the development of verb inflections in child language', in: *Language* 56, 386-412.
- Borer, H. (1984) *Parametric Syntax: Case Studies in Semitic and Romance Languages*, Dordrecht: Foris.
- Borer, H. (2000) 'Exo-skeletal versus Endo-skeletal explanations: Syntactic projections and the lexicon', ms. University of Southern California.
- Braine, M.D.S. (1963) 'The ontogeny of English phrase structure: the first phase', in: *Language* 39, 1-13.
- Brent, M. (1994) 'Surface cues and robust inference as a basis for the early acquisition of subcategorization frames', in: L. Gleitman and B. Landau (eds.) *The Acquisition of the Lexicon*, 433-470, Cambridge Mass./London: MIT Press.
- Bronckart, J.P. and H. Sinclair (1973) 'Time, tense and aspect', in: *Cognition* 2-1, 107-130.
- Brown, R. (1973) *A First Language: the Early Stages*, Cambridge Mass.: Harvard University Press.
- Buszkowski, W. (1987) 'Discovery procedures for categorial grammars', in: E. Klein and J. van Benthem (eds.) *Categories, Polymorphism and Unification*, University of Amsterdam.
- Buszkowski, W. and G. Penn (1990) 'Categorial grammars determined from linguistic data by unification', in: *Studia Logica* 49, 431-454.
- Chomsky, N. (1995) *The Minimalist Program*, Cambridge Mass./London: MIT Press.
- Chomsky, N. (1999) 'Derivation by Phase', *MITOPL* 18, Dept. of Linguistics, MIT.
- Clark, E.V and C.J. Sengul (1978)**
- Coene, M. and L. Avram (2001) 'Acquisition of auxiliaries and determiners in early child Romanian', *Proceedings of ELA (Early Lexicon Acquisition) 2001*, cd-rom.
- Crystal, D. (1987) *The Cambridge Encyclopedia of Language*, Cambridge: Cambridge University press.
- Drozd K.F. (1995) 'Child English pre-sentential negation as meta-linguistic exclamatory sentence negation', in: *Journal of Child Language* 22: 583-610.
- Evers-Vermeul, J. (2000) 'De complexiteit van connectiefverwerving (The complexity of the acquisition of connectives)', in: *Nederlandse Taalkunde* 5-3, 250-271.
- Evers, A. and J. van Kampen (2001) 'E-language, I-language and the order of parameter setting', UiL OTS Working Papers 00105-S-S, Utrecht University.
- Fletcher, P. (1985) *A Child's Learning of English*, Oxford/London: Blackwell.

- Gagarina, N. (2001) 'Aspectuality and its acquisition in Russian', Poster presented at the Perspectives on Aspect conference, December 13, UiL OTS, Utrecht University.
- Gleitman, L. (1990) 'The structural source of verb meaning', in: *Language Acquisition* 1, 3-55.
- Gropen, J. S. Pinker, M. Hollander and R. Goldberg (1991) 'Affectedness and direct objects: the role of lexical semantics in the acquisition of verb-argument structure', in: *Cognition* 41, 153-195.
- Gruber, J. (1975) 'Performative-constative transition in child language development', in: *Foundations of Language* 12, 513-537.
- Haan, G. de (1987) 'A theory-bound approach to the acquisition of verb placement in Dutch', in: G. de Haan and W. Zonneveld (eds.) *Formal Parameters of Generative Grammar* III, 15-30, Yearbook 1987, Utrecht: OTS.
- Hale, K. and S.J. Keyser (1993) 'On argument structure and the lexical expression of syntactic relations', in: K. Hale and S.J. Keyser (eds.) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 53-108, Cambridge Mass: MIT Press.
- Hamann, C. and K. Plunkett (1998) 'Subjectless sentences in child Danish', in: *Cognition* 69, 39-72.
- Heim, I.R. (1982) *The Semantics of Definite and Indefinite Noun Phrases*, PhD. dissertation University of Massachusetts, Amherst.
- Hoekstra, T. and P. Jordens (1994) 'From adjunct to head', in: T. Hoekstra and B.D. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar*, 119-149, Amsterdam/Philadelphia: Benjamins.
- Hoekstra, T. and N. Hyams (1998) 'Aspects of root infinitives', in: *Lingua* 106, 81-112.
- Hulk, A. 'Merging scope-particles. Word order variation and the acquisition of *aussi* and *ook* in a bilingual context', ms. University of Amsterdam.
- Houwer, A. de (1997) 'The role of input in the acquisition of past verb forms in English and Dutch: evidence from a bilingual child' in: E. Clark (ed.) *Proceedings of the 28th Stanford Child Language Research Forum*, Stanford: CSLI, 153-162
- Kampen, J. van (1987) 'De ontwikkeling van ontkenningen in de kindertaal (The development of negation in child language)', ms. Utrecht University.
- Kampen, J. van (1994) 'The learnability of the Left-Branch Condition', in: *Linguistics in the Netherlands* 1994, 83-94.
- Kampen, J. van (1997) *First Steps in Wh-movement*, Delft: Eburon.
- Kampen, J. van (2000a) 'An alternative scenario for the acquisition of syntactic categories', in: *LINK* 12-1, 29-42.
- Kampen, J. van (2000b) 'Left branch extraction as operator movement: Evidence from child Dutch', in: S.M. Powers & C. Hamann (eds.) *The acquisition of scrambling and cliticization*, Kluwer, Norwell, 479-503.
- Kampen, J. van (2001a) 'Review of Becker 'The Development of the Copula in Child English'', in: *GLOT International* 7-5, URL <http://www.linguistlistplus.com>.
- Kampen, J. van (2001b) 'Pragmatic bootstraps for category assignment: from non-discourse towards discourse oriented language', in: Proceedings of ELA (Early Lexicon Acquisition), Lyon, December 5-8, cd-rom.

- Kampen, J. van and F. Wijnen (2000) 'Grammaticale ontwikkeling (Grammatical development)', in: A.M. Schaerlaekens and S. Gillis (eds.) *Kindertaalverwerving*, 225-285, Groningen: Wolters-Noordhoff.
- Karmiloff-Smith, A. (1981) 'The grammatical marking of thematic structure in the development of language production', in: W. Deutsch (ed.) *The child's construction of language*, New York: Academic Press, 121-147.
- Ketrez, F.N. and A. Aksu Koç (2000) 'Emergence of noun and verb categories in Turkish', paper presented at the 10th International Congress on Turkish Linguistics (ICTL X), 16-18 August, Bogazici University, Istanbul, Turkey.
- Krikhaar, E. and F. Wijnen (1995) 'Children's categorization of novel verbs: syntactic cues and semantic bias', in: E. Clark (ed.) *Proceedings of CLRF*, Stanford University.
- Lebeaux, D. (1988) *Language Acquisition and the Form of Grammar*, PhD. dissertation University of Massachusetts.
- Lyons, J. (1977) *Semantics*, Cambridge: Cambridge University Press.
- Lyons, J. (1979) 'Deixis and anaphora', in: T. Myers (ed.) *The development of conversation and discourse*, Edinburgh: University Press, 88-103.
- Maratsos, M. (1982) 'The child's construction of grammatical categories', in: E. Wanner and L.R. Gleitman (eds.), *Language acquisition: the state of the art*, 240-266, Cambridge: Cambridge University Press.
- Maratsos, M. and M. Chalkley (1983) 'The internal language of children's syntax: The ontogenesis and representation of syntactic categories', in: K.E. Nelson (ed.), *Children's language (Vol. 2)*, New York: Gardner Press.
- Marantz, A. (1997). 'No escape from syntax: Don't try morphological analysis in the privacy of your own Lexicon', A. Dimitriadis et.al. (eds.) *Penn Working Papers in Linguistics* 4:2, Proceedings of the 21st Annual Penn Linguistics Colloquium, 201-225.
- O'Grady, W. (1997) *Syntactic Development* Chicago: University of Chicago Press.
- O'Grady, W. (1999) 'The acquisition of syntactic representations', in: W.C. Ritchie and T.K. Bhatia (eds.) *Handbook of Child Language Acquisition*, 157-193, San Diego/London/Boston: Academic Press.
- Penner, Z., R. Tracy and J. Weissenborn (2000) 'Triggering object scrambling', in: S.M. Powers and C. Hamann (eds.) *The Acquisition of Scrambling and Cliticization*, 127-164.
- Pinker, S. (1984) *Language Learnability and Language Development*, Cambridge Mass./London: Harvard University Press.
- Pinker, S. (1987) 'The bootstrapping problem in language acquisition', in: B. MacWhinney (ed.) *Mechanisms of Language Acquisition*, 399-441, London: Lawrence Erlbaum Associates..
- Pinker, S. (1989) *Learnability and Cognition. The Acquisition of Argument Structure*, Cambridge Mass./London: Harvard University Press.
- Pinker, S. (1994) 'How could a child use verb syntax to learn verb semantics?', in: L. Gleitman and B. Landau (eds.) *The Acquisition of the Lexicon*, 377-410, Cambridge Mass.: MIT Press.
- Powers, S. (1996a) *The Growth of the Phrase Marker: Evidence from Subjects*. PhD. dissertation University of Maryland.
- Powers, S. (1996b) 'MAPping phrase markers', in: Ch. Koster and F. Wijnen *Proceedings of GALA* (Groningen Assembly on Language Acquisition), 303-313.

- Powers, S. and D. Lebeaux (1998) 'More data on DP acquisition', in: N. Dittmar and Z. Penner (eds.) *Issues in the Theory of Language Acquisition*, Bern: Peter Lang, 37-76.
- Radford, A. (1996) 'Towards a structure-building model of acquisition', in: H. Clahsen (ed.) *Generative Perspectives on Language Acquisition*, Amsterdam: John Benjamins, 43-89.
- Roeper, T. (1996) 'The role of merger theory and formal features in acquisition', in: H. Clahsen (ed.) *Generative Perspectives on Language Acquisition*. Amsterdam: John Benjamins, 414-449.
- Saussure, F. de (1915) *Cours de Linguistique Générale* Ch. Bally, A. Sechehaye & A. Riedlinger (eds), reprint 1964, Paris: Payot.
- Schönenberger, M., Z. Penner and J. Weissenborn (1997) 'Object placement and early German grammar', in: E. Hughes, M. Hughes, and A. Greenhill (eds.) *Proceedings of the Annual Boston Conference on Language Development 21*, Somerville MA: Cascadilla Press.
- Stassen, L. (1997). *Intransitive Predication*. Oxford Studies in Typology and Linguistic Theory. New York: Oxford University Press.
- Strawson, P.F. (1974) *Subject and Predicate in Logic and Grammar*, London: Methuen.
- Verrips, M. (1996) *Potatoes must peel: the Acquisition of the Dutch Passive* PhD. dissertation University of Amsterdam.
- Wijnen, F. (2000) 'Input, intake and sequence in syntactic development', in: M. Beers, B. van den Bogaerde, G. Bol, J. de Jong and C. Rooijmans (eds.) *Proceedings of the Child Language Seminar 'From sound to sentence'*, Garderen, September 1997, 163-186, Groningen: Center for Language and Cognition.

Appendix A

I. *Examples: Pragmatic constructs with an illocution operator and a category-neutral content sign X.* (Only a sample of the tokens that occur in the Sarah files is given. The numbers refer to the file numbers and *d.n.* to diary notes. The sign -- stands for “doesn’t occur in the files”).

Dutch (Sarah, Van Kampen corpus, till week 150)

	illocution operator	category-neutral X (later N - V - P - A)			
a.	dat }	maan 08	--	uit 08	open 08
b.	dit }	muisje 06	lachen 09	in 06	veel groot 07
c.	die } (ook) (niet)	hertje 06	vallen 01	aan 06	au 04
d.	deze }	eendje 06	huilen 15	op 02	lief 11
e.	hier }	beer 05	wassen 08	boven 08	--
f.	daar }	aap 05	wrijven 12	uit 08	leuk 09
g.	niet	huis 02	tekenen 09	uit 06	stout 02
h.	(ik) mag (wel)	kaart 11	kleuren 09	--	vies d.n.
i.	ik ook	boot 04	hebben 05	mee 12	bruin 13
j.	(i)kwi(l)	varkje 07	tikken 04	af 08	mooi d.n.
k.	(i)k wi(l) niet	vliegje 13	sturen 14	uit 11	--
l.	(ik) hoef (niet/wel)	hondje 09	wassen 09	--	--
m.	(ik) kan (niet/wel)	neus 11	zien 11	bij 13	hoog 26
n.	nou }	Petteflet 15	slapen 13	uit 09	donker 15
o.	ook } (niet)	paard 03	puzzelen 07	buiten 08	leuk 17
p.	nog }	aap 05	zoeken 09	--	groot 21
q.	even	schuitje 07	zitten 08	d'r uit 19	klaar 19

a.	that	moon/--/out/open	j.	(I) wanna	piglet/tap/off/pretty
b.	this	mouse/laugh/in/too big	k.	(I) wanna not	fly/send/out/--
c.	that	deer/fall/on/ouch	l.	(I) need not	doggy/wash/--/--
d.	this	duck/cry/on/sweet	m.	(I) cannot	nose/see/with/high
e.	here	bear/wash/above/--	n.	now	Petteflet/sleep/out/dark
f.	there	monkey/rub/out/nice	o.	also	horse/(to) puzzle/outside/nice
g.	not	house/draw/out/naughty	p.	yet (=more)	monkey/search/--/big
h.	I may	card/color/--/dirty	q.	just	boat/sit/out/ready
i.	I also	boat/have/with/brown			

II. *Examples: Pragmatic constructs with a negation operator and a category-neutral X.*

Dutch (Tim, 2;1-2;4, Trum corpus; Van Kampen 1987)

illocution operator	category-neutral X (later N - V - P - A)			
<i>deictic neg.</i> (is) niet ((is) not)	hondje (doggie) de botsing (the crash) mama (mummy)	slapen (sleep) opleggen (lay on)	in (in)	kapot (broken) au (ouch)
<i>deontic neg.</i> (kan/mag) niet ((can/may) not)	koek (cake) bootje (boat)	rijden (drive) uitkijken (look out) opendoen (open) eten (eat) huilen (cry) kapot maken (destroy) pikken (pinch) zien (see) horen goed (hear good) draaien (turn)	aan (on) op (at) af (off) langs (along)	open (open)

Appendix B

Dutch (Van Kampen corpus)

I. Numbers: Occurrences of functional words used as pragmatic constants

(Sarah, week 86-150)

Illocution operator	Number	Adult language	
<i>deictic operator</i>			
dat (is) X	118	} demonstrative	(that (is) X)
dit (is) X	162		(this (is) X)
die (is) X	160	}	(that (is) X)
deze (is) X	123		(this (is) X)
hier (is) X	99	} place adverb	(here (is) X)
daar (is) X	49		(there (is) X)
niet (is) X	102	negation adverb	(not (is) X)
<i>modal operator</i>			
(ik/jij) kan (wel) X	100	deontic modal	((I) can (indeed) X)
(ik) mag (wel) X	132	deontic modal	((I) may (indeed) X)
ik ook X }	141	} first person pronoun subject	(I also X)
ik ∅ X }			} + cliticized modal
(i)kwi(l) X }		}	(I) wanna X)
ik/ik ∅ niet X }	144		} first person pronoun subject
(ik) wil niet X }		} + volitional modal	(I want not X)
(ik) kan niet X }		} + negation adverb	(I cannot X)
(ik) hoef niet X }		}	(I need not X)
<i>turn taker</i>			
nou X	140	} pragmatic sentence adverb	(now X)
nog X	105		} c.q temporal adverb
ook X	183		
even X	66		(just X)
zo X	49		(so X)
<i>place pointer</i>			
X nou ?	166	} pragmatic sentence adverb	(X now?)
X daar	98		} place adverb
X hier	37		
X ook	96	adverb of addition	(X also)

II. Numbers: All occurrences of functional words (pragmatic constants and adult use)

(Sarah, week 86-150)

pronoun/adverb		verb		adverb	
dat	303	(ik) wil	208	wel	100
deze	383	kan	147	niet	719
die	373	mag	136	ook	434
dit	345	moet	183	nog	194
daar	154	ga(at)	244	nou	286
hier	136			even	158
				zo	94

Appendix C

A. Presentationals with the deictic operator referring to an entire situation

Dutch (Laura S. 3;0, Schaeffer p.c.; some examples; see Van Kampen 1997: 126f)

a. dat	slordig opeten	(that (is) (to) eat careless)
b. dat	peer slordig opeten	(that (is) (to) pear eat careless)
c. dat	goed opeten	(that (is) (to) eat well)
d. dat	lang(zaam) opeten	(that (is) (to) eat slowly)
e. dat	goed uitknippen	(that (is) (to) cut out well)
f. dat	lang(zaam) uitknippen	(that (is) (to) cut out slowly)
g. dat	lelijk doen	(that (is) (to) do nasty)
h. dat	mooi doen	(that (is) (to) do beautifully)
i. dat	vlug natekenen	(that (is) (to) copy quickly)

Dutch (Laura, Van Kampen corpus)

a. die	zo water in .	(2;6.10)
b. hier	papa autorijden	(2;8.24)
c. dat isse	papa autorijden	(2;8.24)
d. dat is	Wim autorijden	(2;8.24)

English (Nina, Suppes corpus)

a. thats	holding a balloon	(1;11.29)
b. thats	my fish drinking a water	(1;11.29)
c. thats	horsie running	(2;03)
d. thats	Mommy holding a baby	(2;0.24)
e. that	teddy bears on it	(2;0.24)
f. thats	him been in my house	(2;1.6)
g. thats	door white	(2;1.15)
h. that	pink in that bedroom	(2;1.15)
i. that	top in food	(2;1.15)
j. that	the door close	(2;1.15)
k. thats	playing the doll house	(2;1.22)
l. its	a ribbon on the dog	(2;2.12)
m. thats	a don't move	(2;2.12)
n. thats	a turtle swim	(2;2.28)
o. thats	turn around there see the train	(2;2.28)

English (Braine 1976, Powers 1996a)

f. here	more book/milk	(David 1;10)
f. there	bye bye car	(Gregory)
g. there	daddy truck	(Gregory)
h. that	daddy there	(Daniel 1;8)

Swiss German (examples from Penner 1993, Käsermann corpus; see Van Kampen 1997:126f)

a. das	nid Selleri gärn	(Simone 2;0)	(that not Selleri like)
b. das	ou Loki ditue	(Reto 2;0)	(that also locomotive put inside)

French (Philippe, Leveillé corpus; possible cases)

a. c' est Myriam a cassé le tracteur	(2;1.19)	(it's Myriam has broken the tractor)
b. c' est des boutons	(2;1.19)	(it's buttons)
c. c' est trop dur faire la maison	(2;1.26)	(it's too hard make the house)
d. c' est des garçons va à l' école	(2;2.3)	(it's boys go to school)
e. c' est la maîtresse a des crayons et des stylos	(2;7.25)	(it's the teacher has pencils and pens)
f. c' est des gâteaux elle a porté Madeleine	(2;8.1)	(it's cookies she has brought M.)
g. c' est des papiers pour travailler	(2;8.22)	(it's papers for working)
h. c' est les chinois ils font ça	(2;11.7)	(it's the Chinese they do it)
i. c' est des voitures avec deux portes ouvertes	(2;11.7)	(it's cars with two open doors)
j. c' est tous des camions ça	(2;11.7)	(it's all lorries that)

Presentationals with the copula as an utterance announcer

(Laura , Van Kampen corpus)

HEEL VEEL

(Adam, Brown corpus, 11-16)

a. is: cowboy	11
b. is: clown	11
c. is: a pipe	11
d. is: (de)licious	12
e. is: good	12
f. is: nice	14
g. is: not a toy	14
h. is: goes there	16
i. is: going with airplane	16
j. is: walks	16

B. Modal operator referring to an entire situation

Dutch (Sarah, Van Kampen corpus)

a. ook	schoen uit	(1;11.15)	(wanna shoe off)
b. (i)kwi(l)	open	(2;0.17)	(wanna (bottle) open)
c. ikke wi(l)	dit mooi	(2;0.17)	(wanna this beautiful)
d. (i)kwi(l)	(ander)som	(2;1.10)	(wanna (book) other way round)
e. (i)kwi(l)	deze hier	(2;1.10)	(wanna this here)
f. (ik) moet	(ander)som	(2;1.10)	((I) must (book) other way round)
g. mag niet	<ha, ha,ha>	(2;0.22, diary)	((I) may not do <ha, ha, ha>)
h. moet	liedje aan	(2;3.13, diary)	((I) must (turn) song on)

English (examples from Radford 1990/1996)

a. wanna	baby talking	(Haley 1;8)
b. wanna	mummy come	(Jem 1;9)
c. wanna	mummy take it out	(Jem 2;0)
d. wanna	this go up	(Angharad 1;10)
e. wanna	teddy drink	(Daniel 1;7)
f. wanna	lady open it	(Daniel 1;10)
g. wanna	lady get chocolate	(Daniel 1;11)
h. wanna	mummy do	(Anna 2;0)
i. wanna	chair put in	(Lucy 2;0)

BOWERMAN checken

C. Turn taker referring to an entire situation

Dutch (Sarah, van Kampen corpus)

a. ook	ete(n) t(r)ein	(1;11.15)	(also ∂ (to) eat (in) train)
b. ook	beer in bad	(1;11.15)	(also (is) bear in bath)
c. ook	allemaal kip	(1;11.15)	(also (is) all chicken)
d. ook	bloe(d) op ∂ duim	(2;2.18)	(also (is) blood on ∂ leg)
e. nog ∂	spullen weg	(2;0.17)	(still ∂ (are) things gone)
f. nou	allemaal kapot	(2;2.18)	(now (is) all broken)
g. even	schuitje	(1;11.1)	(just (go) boat(ing) = a children's game)

D. Negation operator referring to an entire situation

I. Presentational negation

Dutch (Laura S. 3;0, Schaeffer p.c.; some examples)

a. dat	niet	banaan opeten	(that (is) not (to) eat banana)
b. dat	niet	aap uitknippen	(that (is) not (to) cut out monkey)
c. dat	niet	peertje schillen	(that (is) not (to) peel pear)
d. dat	niet	snoepje opeten	(that (is) not (to) eat candy)
e. dat	niet	boterham opeten	(that (is) not (to) eat sandwich)
f.	niet	boterham eten	(that is) not (to) eat sandwich)
g.	niet	de hond uitknippen	((that is) not (to) cut out doggie)
i.	niet	twee doen	(that is) not (to) make two)

Dutch (Sarah, van Kampen corpus)

a. dit is	niet	bad zwemmen.	(1;11.15)	(this is not (to) bath swim)
b. dit ∂	niet	paard zwemmen	(1;11.15)	(this ∂ not horse swim(ming))
b.	niet ∂	bad zwemmen	(1;11.15)	((this is) not (to) bath swim)
c.	niet	∂ paard zwemmen	(1;11.15)	((this is) not ∂ horse swim(ming))
d. is	niet	teken(en)!	(2;0.17)	((this) is not (to) draw)
e.	niet ∂	teken(en)	(2;0.17)	((this is) not ∂ (to) draw)
f.	niet ∂	lachen.	(2;0.17)	((this is) not ∂ (to) laugh)
g. die	niet	lachen	(2;0.17)	(that (is) not (to) laugh)
h. dit ∂	niet	lachen	(2;0.17)	(this ∂ (is) not (to) laugh)
i. dit ∂ is	niet	lachen	(2;0.17)	(this ∂ is not (to) laugh)

to be continued