

The Learnability of Syntactic Categories

Jacqueline van Kampen
UiL OTS, Utrecht University

1. UG as the outcome of the acquisition process

It is a common position in generative acquisition studies to accept Chomsky's view that first language acquisition is determined by a set of innate grammatical priories. The development of the child would be more a matter of biological maturation than a matter of input-control. Because language universals are innate in the human mind, they cause grammar to grow into the mind almost automatically under the slightest provocations. Early child language would already show the relevance the grammatical priories. Generative grammarians guided by this view have often drawn far-reaching conclusions about the structure of early child language. The present paper will present an alternative view, the derivation of UG principles from structural acquisition steps. It acknowledges that it is indeed a sentence-generating system that is acquired, but contends that generative systems are learned from the language-specific input material. The basic argument for this approach is that all eventual 'UG' properties are identified due to local relations with language specific shapes. One might see the language specific shapes as an entrance to the UG distinctions. Unless a grammar offers a way to identify UG properties, it will not be learnable. This suggests that UG properties may be seen as the outcome of an acquisition procedure rather than being its source.

The possibility and the effectiveness of the stepwise procedure follow from a fundamental property of the grammatical system, its linear locality. When one looks at the simplified structures of child language, the following points are relevant. Initially, the child has little else to adhere to, since she has no grammar yet as she is still on her way to find out. Rules for stress assignment, phrasal and word categories, - lexical or grammatical, abstract or empty -, cannot be applied until there is at least some grammatical environment. It does not matter that much whether the child is or is not informed a priori about the structure of grammar. Such knowledge is largely ineffective anyway. There has to be some search for a primary language specific orientation first. Fortunately, the stepwise progress of that search can be followed by the longitudinal analysis of the child's development. The longitudinal analysis shows how the child adds each time grammatical markers within a preceding binary locality frame. The two members of a frame have no internal structure in early child language. Later child language maintains the orientation by binary frames, although by then the two members of the frame may themselves have internal structure. Each syntactic acquisition step relies on such a preceding binary frame. These preliminary acquisition frames – called *evidence frames* (Van Kampen 1997) – remain present in adult language as checking domains. In this view, checking domains testify a prehistory in which they were the origin of an acquisition step. I would like to present the orientation on the order of the learning steps as the *learnability research*. The learnability research implies an insistence that natural grammars are 'designed' to be learned by children. The learnability perspective holds for language-specific arrangements as well as for grammatical universals.

1.1 Outline of the paper

I cannot immediately provide reinterpretations for all arguments in the literature that UG conditions must be innate, but this paper will argue the 3 acquisition steps below. The argumentation will be based on examples from English and Dutch child data in CHILDES (MacWhinney 1991).

1. Proto-grammar. There is some initial grammar before the acquisition of finite verbs and determiners. Following a suggestion by Lyons (1979), I will indicate that initial grammar as proto-grammar Proto-grammar is based on illocution operators for simple pragmatic language games {*more, no, want, is, that, there, where, what*} and lexical content elements {*bear, drink, car, nice*}. I will exemplify that distinction in *section 2*. The claims I make are in line with the analyses by Lebeaux (1988), Radford (1990) and Jordens (2002), but I add the perspective that UG properties are the outcome rather than the source of structural acquisition steps. Elements like wh-words {*what, where*} and finite ‘verbs’ {*wanna, is*} do not necessarily need deep properties from UG syntax in very early child language. They have in the proto-grammar all properties of the context-free operators. The lexical content elements are in the proto-grammar category-neutral between N/V/A/P.

2. The illocution operators develop into functional categories. The child starts the acquisition of adult grammar with predicate marking (I⁰/Aux) versus argument marking (D⁰/Case). The source of that development is a grammatical marking of topic-comment intentions as IP structures. For example *bear running/nice* becomes *bear is running/nice*. The illocution operators in proto-grammar are highly situation-related whereas the functional categories like the various types of I⁰ and D⁰ are by contrast highly sensitive to syntactic context. The first functional category that is acquired, at least in Dutch, French and English, is I-marking, the marking of a predicate by a factor <+I> {copula, auxiliary, modal, finite morphology (*section 3*).¹ I-marking is followed by the acquisition of D-marking, the marking of arguments by a factor <+D> {article, demonstrative, possessor, quantifier} (*section 4*).

3. The lexical content elements are labelled N/V/A/P. The acquisition of systematic I/D-marking is the key to the acquisition of category distinction between V/N (*sections 3-4*). The universal lexical categories <+N> and <+V> are acquired from the language specific D- and I-marking. Note how this presumes an acquisition path from a syntactic distinction (IP/DP) towards the lexical category distinction (V/N), playing down the cognitive distinction between events/things. It is a reversal of Pinker’s bootstrapping scenario, which moves from a cognitive distinction (event/thing) towards the lexical classes (V/N) before arriving in syntax (IP/DP).

2. Proto-grammar

There is a period in child language when utterances are mostly binary sign combinations. I use the Saussurian word ‘sign’ here as a hedge, since quite a few of these elements will be reanalyzed later on as morphological or phrasal constructs: *wanna* becomes *want to* or *want a*, and for Dutch *kannie* (‘cannot’) becomes *kan niet*. See Bellugi (1967), V.Kampen (2001), and Jordens (2002) for more examples.

¹ An account of the <-fin> utterances in early child language should include not only later verbal predicates, so-called Optional ‘Infinitives’, but also later non-verbal predicates like ‘daddy nice’ or ‘bear in (the) zoo’. See Van Kampen (1997) for this generalization.

The content elements in this period do not carry by themselves a referential or a predicative intention, nor are they marked by some I° for predication or some D° for reference. I assume this is the period that Lyons (1979:90) had in mind when he suggested that child language might have proto-predication as a forerunner of predication and proto-reference as a forerunner of reference. I propose to give more grammatical and empirical substance to these ideas.

- | | | | |
|-----|---|---|--|
| (1) | <u>Proto-grammar</u> | | <u>Grammar</u> |
| a. | proto-reference
“topic naming” | → | reference (systematic D-marking)
D° + X |
| b. | proto-predication
“characterizing comment” | → | predication (systematic I-marking)
I° + X |

Let predication and reference be distinctions that can be defined only in connection with a grammar. Early child language may use a sign to characterize some salient aspect of the situation. Let me call that proto-predication rather than predication. Let the notion predication be applied only if the characterizing signs are systematically marked by one of the I° devices {copula/modal/auxiliary/finite morphology}. The same recipe applies to reference. Early child language may use a sign to name a salient aspect of the situation. Let me call that proto-reference rather than reference. Let the notion reference be applied only if the naming signs are systematically marked by one of the D° devices {article/possessive/demonstrative/ case marking}. The functional categories I° and D° are second order elements. They create and mark a syntactic frame for content elements.

2.1 Category-neutral content signs X

Bear in (4) may be regularly used by the child without a copula I° or a marker D°. It may be used as a situation-bound topic and proper name (proto-reference) or as a situation-bound characterizing comment (proto-predication).² The switch between naming and characterizing signs is not yet a matter of syntactic categories but a matter of pragmatic intention only, see (2).

- | | | |
|-----|---|-------------------------|
| (2) | Pragmatic switch of intention for category neutral X | |
| a. | proto-reference | b. proto-predication |
| | | |
| | X | X |
| | <i>beertje</i> (bear) | <i>beertje</i> (bear) |
| | (is a proper name with fixed reference,
i.e. name for the child's cuddly bear) | (is a characterization) |

The category neutral X° survives in binary utterances with two content signs (Van Kampen 1997, 2001), see the examples in (3).³

² I make a distinction between the situational versus the (linguistic) discourse context of a sentence. The situational context is in principle available to the child without additional grammatical devices. The discourse context is in not accessible to the child until there is the systematic application of grammatical deictic devices D° and I° (cf. Hoekstra & Hyams 1998).

³ The analysis of category-neutral binary structures in child language was elaborated in Van Kampen (1997) for the Y-X° pattern in (5) and X°-Y patterns like *read booklet/from daddy*.

(3) Topic-comment structures (two category-neutral content signs)			
<u>Dutch</u>		<u>English</u>	
XP proto-predication		XP proto-predication	
proto-ref	proto-pred	proto-ref	proto-pred
Y	X ^o	Y	X ^o
<i>beertje</i>	<i>slapen</i>	<i>bunny</i>	<i>dance(-ing)</i>
<i>mamma</i>	<i>dom</i>	<i>girl</i>	<i>bad</i>
<i>aapje</i>	<i>op</i>	<i>shoes</i>	<i>off</i>
<i>pop</i>	<i>meisje</i>	<i>doggie</i>	<i>girl</i>

I characterize (3) as topic-comment structures that function as proto-predications. The utterance stress is usually on the second element, the proto-predicate. The two-word utterances are in the present view category-neutral. Lexical categories <+V>/<+N> cannot be assigned before there is an idiomatic relation with a predicate marker I^o or a reference marker D^o. The same reserve is recommendable for the terms subject and predicate. Children may have sensible reactions, but interpreting them by means of formal grammatical understanding is premature. The subject-predicate relation follows from the form-based I^o/D^o devices of real grammar. These are atypical for proto-language. The proto-predicate X^o is not systematically marked by some I^o, whereas the later predicate is. Systematic non-expression of a distinction is best explained by systematic absence of the distinction.

2.2 Pragmatic illocution operators F

Category neutrality also holds for another type of binary utterances. This type combines a category-neutral content sign with a pragmatic operator F^o (Van Kampen 2001).⁴ The operator marks the utterance for some fixed use in a language game, see (4). There is a presentational/naming in (4)a, a wish in (4)b, a sign for conversational turn taking in (4)c, a denial in (4)d, a refusal in (4)e, a possession claim in (4)f and a content question in (5). The operator appears either in left-handed position (operator+X (4)) or in right-handed position (X+operator (5)a).

(4) Illocution-marked proto-predications (operator + category-neutral content sign)

	FP<+illoc>		FP<+illoc>		
	F ^o (operator)	X (proto-pred)	F ^o (operator)	X (proto-pred)	F ^o is
a.	<i>is</i>	<i>muisje/lief/spelen</i>	<i>is</i>	<i>mouse/nice/play</i>	operator for: presentation naming wish turn-taking denial refusal possession claim
	<i>dat</i>	<i>muisje</i>	<i>that</i>	<i>mouse</i>	
b.	<i>kwil</i>	<i>spelen</i>	<i>wanna</i>	<i>play</i>	
c.	<i>nog/ook</i>	<i>op</i>	<i>more</i>	<i>on</i>	
d.	<i>niet</i>	<i>lief</i>	<i>not</i>	<i>nice</i>	
e.	<i>nee</i>	<i>beer</i>	<i>no</i>	<i>bear</i>	
f.	<i>ikke</i>	<i>muis</i>	<i>my</i>	<i>mouse</i>	

⁴ The use of a category X is not uncommon as an expository notation. E.g. the operators *wanna* and *more/no* {P:*up*; V:*see*; N:*car*; A:*dirty*} in child language are characterized as *want X* and *more/no X* in Powers (1996) and Powers & Lebeaux (1998). It is proposed here and in Van Kampen (1997, 2001) that the notation be taken seriously, and that the categories <+/-N>, <+/-V> be derived later on from discriminating functional contexts.

- (5) a. $\text{FP}<+\text{illoc}>$
 $\swarrow \quad \searrow$
 X (proto-pred) F^o (operator)
muisje *nou ?*
- b. $\text{FP}<+\text{illoc}>$
 $\swarrow \quad \searrow$
 F^o (operator) X(proto-pred) operator for:
where *mouse?* content question

See for examples also Braine (1963), who labeled them *pivots*, Roeper (1992) and Jordens (2002), who presents a similar operator analysis for early child Dutch. None of these studies propose category-neutrality, but I do.

The illocution operators introduce a comment sign. The same is present in the examples in (6)-(7) where the operator functions as an utterance announcer and refers to an entire situation. See for more examples Van Kampen (2001), Jordens (2002).

- (6) Dutch Sarah (Van Kampen corpus)
- | | | | | |
|----|---------|-------------|--------------------|-----------------------------|
| a. | dit∂: | hondje weg | (1;10.5, week 96) | (this∂ (is) doggie gone) |
| b. | kwi(1): | dit mooi | (2;0.17, week 107) | (wanna this beautiful) |
| c. | kan: | deur | (2;5.22, week 129) | ((I) can (open the) door) |
| d. | manie: | <ha, ha,ha> | (2;0.22, week 108) | ((I) may not (do) <ha, ha>) |
| e. | nog: | paardje | (2;1.10, week 110) | (‘more’ (you play) horsie) |
- (7) English child language
- | | | | | |
|----|---------|----------------|---------------|---------------------|
| a. | that: | daddy there | (Daniel 1;8) | (from Braine 1963) |
| b. | that’s: | horsie running | (Nina 2;0.3) | (Suppes corpus) |
| c. | is∂: | goes there | (Adam 2;10) | (Brown corpus) |
| d. | wanna: | lady open it | (Daniel 1;10) | (from Radford 1996) |
| e. | no: | lamb have it | (Nina 2;0.24) | (Suppes corpus) |

According to my counts in proto-grammar (8), the operators tend to have a frequency that is a few hundred times higher than the content signs. I propose that the learning device is sensitive to frequency difference.

- (8) Number of occurrences for single items (types) (Dutch Sarah, week 86-150)
- | | | |
|----|-------------------------------|-----------------------|
| a. | <i>Illocution operators</i> : | between 50-300 tokens |
| b. | <i>Content signs X</i> : | between 1-30 tokens |

The pragmatic operators are not a kind of functional categories. They can only be defined by means of situation types, whereas functional categories such as articles and copulas can only be defined within syntax. The introduction of the pragmatic operators does not imply either that the abstract syntax of the common generative approach has now been replaced by complex pragmatics. The distinctions of the pragmatic operators are elementary situation-bound oppositions between roles in elementary games (Wittgenstein 1953).

There is something remarkable about the presentational operators in (4)a,b and the question operators in (5). The presentationals (4)a,b are taken from a demonstrative *that/this* as well as from a copula *is*. These very signs in proto-grammar have a singularizing pointing effect. At a certain point in acquisition, they are added regularly to content signs for topic and comment. As soon as the comment function is systematically marked by Aux/I^o and the naming intention by D^o/case, there is a form-based marking of predication and argument structure. The early wh-element in (5)b is arguably an operator that solicits for some verbal reaction of the

conversation partner. Rather than using the *wh*-element, Dutch proto-grammar uses a sentential adverbial (*nou*) (5)a. The *wh*-pronoun is spelled-out when D-marking is acquired. The difference between early content questions in English and Dutch is explained if we assume a simple difference in proto-grammar. A stressed sentence adverbial, not present in English, is generalized into a question operator in (5)a.

The next two sections will follow the development from proto-grammar to real grammar for Dutch Sarah (CHILDES, Van Kampen corpus). At the proto-grammar stage, when her utterances are mainly binary sign combinations Sarah is 1;9.10 (week 93). A year later, at 2;9.7 (week 145), she has acquired I- and D-marking.

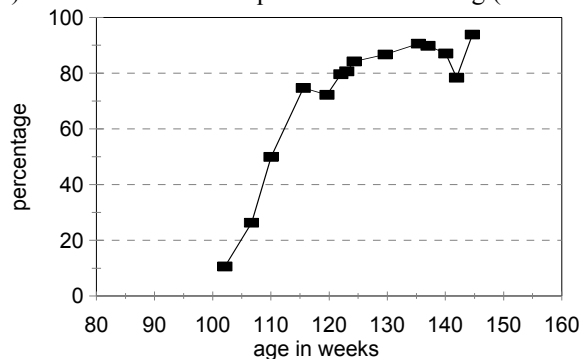
3. The acquisition of <+I> and subsequently <+V>.

From the beginning on, Dutch Sarah had some part of the proto-predicates (>10%) preceded by a clitic-like *-s* that reminds of the copula *is*. Also marginally, proto-predicates began with a modal or finite verb, instead of the unmarked root infinitive.

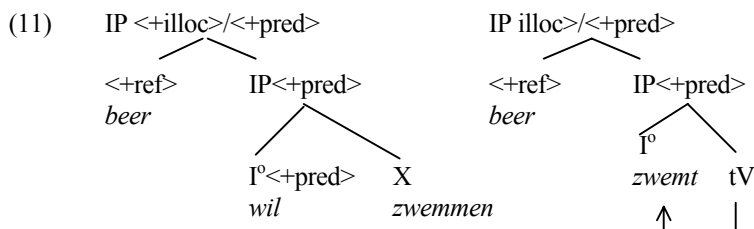
- (9) a. *bal weg* (ball away) *bal ook weg* (ball also away) *bal rollen* (ball roll)
 b. *bal sweg* (ball saway) *bal moet weg* (ball must (go) away)

The type of predicate marking exemplified in the (9)b variants lingers around from the beginning on, but in a marginal way only. From a certain point on, predicate I-marking begins a steep and irreversible rise. Within twenty weeks (weeks 100-120) Sarah's <+fin> predications rise from 10% to more than 90%, see the acquisition graph in (10).

(10) Dutch Sarah: Acquisition of I-marking (from Evers & Van Kampen 2001)



The rise of the graph for I-marking in (10) follows from two separate developments. Modal operators and the copula appear increasingly as prefixes of comments in the topic-comment utterances. This is followed by the rise of finite lexical verbs, see (11).



The V<+fin> movement in V-second Dutch is C° attracted. That problem is neglected in (11), but see Evers & Van Kampen (2001) for a discussion. The rise in I° marking reflected in the graph is part of a larger phrasal development in (12) and (13). The development in (12) represents the acquisition of the subject-predicate construction.

- (12) a. The proto-predicates get systematically marked by I°.
 b. Most I-marked comments are accompanied by topics that have a fixed theta-relation {agent, theme}.

The acquisition procedure perceives the obligatory presence of such topics and establishes the EPP, that is all predicates need a subject, as stated in (13).

- (13) Modal verbs are used as I-markers that need a subject. The use of modals as operators, quite common in proto-grammar, gets marginalized.

A clear indication that the modal operators get blocked by the <+fin> interpretation is the first appearance of the personal pronouns *ik/jij* (I/you), see (14)b below. In the proto-grammar *ik/jij* do not appear consistently, but they are often present as subjects ‘implied’ by the operator, see Van Kampen (to appear). The operator *wil/kwil* (‘wanna’) in ((14)a) meant *ik wil* (‘I want’) and *hoenie* meant *ik hoef niet* (‘I don’t have to’) (Van Kampen 1997, Powers 1996). The same delay was found for the acquisition of *je/tu* in French (Van Kampen to appear). As soon as the <+fin> interpretation of the modal overrules its former use as operator, the EPP comes in and there is a need to fill up the Spec-position. The 1st and 2nd person are now understood and used as a kind of topic/referent. Soon, the personal pronouns are no longer mode-implied elements only. They appear as regular argument options with all predicates. The grammatical situation that results from (12) and (13) allows the constructions to be reinterpreted as the subject-predicate pattern in (14)c.

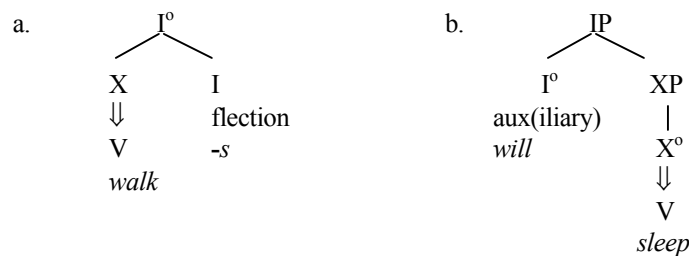
- (14) Dutch Sarah
- | | | | |
|----|--|------------------------------|---------------------|
| a. | <i>First person operator-implied</i> | | |
| | <i>kwi</i> (I) uit | (wanna out) | (1;11.15, week 102) |
| | kan liedje niet | ((I) can song not (sing)) | (2;0.17, week 107) |
| | hoef∂-niet wassen | ((I) need∂ not wash) | (2;0.17, week 107) |
| | kga-even kleuren | (I-go just color) | (2;1.10, week 110) |
| | mag-wel vies | ((I) may indeed dirty) | (2;0.10 diary) |
| b. | <i>First person expressed due to EPP</i> | | |
| | ik wil jou niet spelen | (I want (with) you not play) | (2;1.10, week 110) |
| | ik kan die lezen | (I can that read) | (2;4.27, week 125) |
| | ik hoef niet | (I need not) | (2;4.2, week 122) |
| c. | <i>Modal as independent predicate</i> | | |
| | ik hoef | (I need) | (2;1.10, week 110) |
| | ik wil | (I want) | (2;2.18, week 116) |
| | ik kan niet | (I cannot) | (2;4.27, week 125) |

The common source of the phenomena in (12) and (13) is the reinterpretation of a pragmatic sign for utterance modes as a functional category I° for predicative intentions, see (15). The term ‘bootstrapping’ was introduced by Pinker (1984) for the acquisition of category labels due to cognitive distinctions (thing/event). Here the bootstrap is offered by syntax itself.

- (15) *Bootstrapping* I°
 Pragmatic intention of commenting a topic
 ↓
 Comment optionally expressed by a gesture sustainable deictic operator F°
 ↓
 Obligatory presence of F° equals I° (predication) and EPP

The interpretation of the pragmatic intention <+pred> as <+I> has still another consequence. The <+I> stands for a hand of various devices, some morphological like <+fin> others syntactical as the copula and the other auxiliary-like elements, see (16). Not all elements of the category neutral X can be combined with all of the I°'s. The Dutch copula can be combined with all, but other I-markings are more selective. The three I° types 'X°<+fin>' *is aan het X°* (related to the English present continuous) and *is/heeft X°* <past participle> (related to the English perfect) require a subset of X° elements that scholars may recognize as <+V>. The rules in (17) offer the acquisition context for it (cf. also Blom and Krikhaar 2002).

- (16) lexical <+V> due to morphological context or lexical <+V> due to phrasal context



- (17) Dutch
 X<lex> ⇒ <+V> / [(is) *aan het* —] ('is on the' infinitive)
 / [heeft/is *ge-* —] (past part. morpheme)
 / [<+fin>] (finite verb)

The context sensitive acquisition rules in (16) boil down to the traditional position that V° is identified by its paradigm, where paradigm generalizes over periphrastic and inflectional forms. Slightly different, one might say that the category <+V> is identified by the I° categories in its extended projection (Grimshaw 1991). See for a recent analysis of defining N and V in the syntax Marantz (1997). The new perspective here and in Van Kampen (1997, 2001) is that this common approach to defining V° is seen as a bootstrapping acquisition procedure.

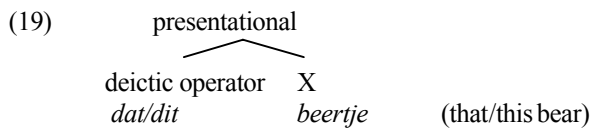
Usually, psycholinguists like cognitive distinctions to get control over grammatical arrangements. As such, it must be attractive to accept Pinker's (1984) claim that the cognitive distinction between things/events lead to lexical categories N/V. Let me point out that the functional categories I and D have good points as well to be considered as bootstraps, see (18) (cf. Emonds 1985:191).

- (18) Functional categories F° differ from lexical content elements X° by
 a. a fixed distribution (order and stress)
 b. a high, closed class, frequency

A good point of the present I/D entrance to lexical categories seems to me that the system itself is analyzed as ‘learner friendly’.

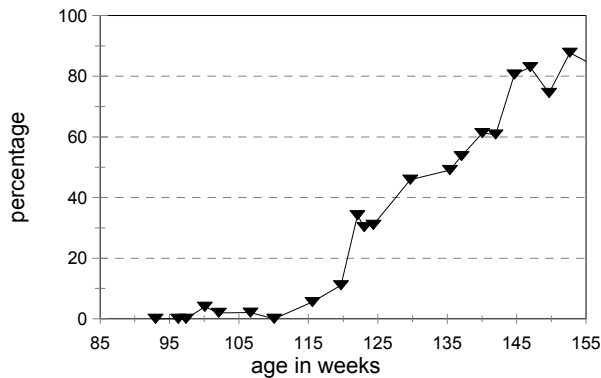
4. The acquisition of <+D> and subsequently <+N>

Consider again the constructs in (19) with an illocution operator and a content element. The deictic demonstratives *dat/dit* (‘that/this’) for the pointing gesture in presentationals have at first an intonation break. They cannot be used in a predicative way and are exclusively referential. In ‘picture looking’ games, or somewhat freer ‘naming games’ they refer to a thing or person.

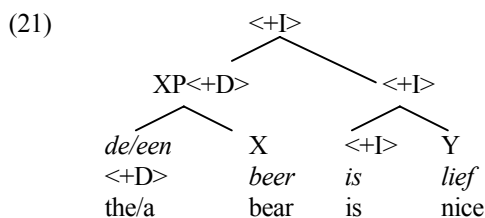


Content signs with a naming function begin to appear with articles in a fashion strikingly similar to the I-marking of the proto-predicates, but 20 weeks later. The insertion of articles (D⁰) rises again from less than 10% to more than 80% within a period of 20 weeks. See the acquisition graph in (20) (from Van Kampen 2001).

(20) Dutch Sarah: Acquisition of D-marking of arguments



The rise of the graph for D-marking in (20) illustrates how articles *de/een* (‘the/a’) appear as phrasal ‘prefixes’ of a referentially intended X (first opposition <+/-definite>) and begin the irreversible rise at week 120. An example of the introduction of the article is given in (21).



The rise of D-marking reflected in (20) is part of the larger development in (22).

- (22) a. The proto-reference part of the utterance is marked by D⁰ (demonstrat., article)
 b. The deictic operator for naming things *dat* changes into an independent topic/subject demonstrative due to the EPP.
 c. The demonstrative *dat* and the 3rd person pronouns appear in all argument positions. Reference marking is perceived and learned as the hallmark of argument-hood.

Examples of (22)a, (22)b and (22)c are given in (23)a, (23)b and (23)c respectively.

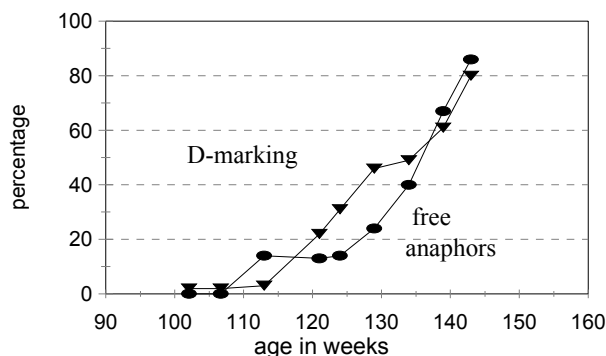
- (23) a. [ik wil [*deze* spelletje]_{DP}] (I want this game) (Sarah 2;3.16, week 120).
 b. *dat* is een paddestoel (that is a mushroom) (Sarah 2;3.16, week 120).
 c. (Mother): is die 'afgelopen', het boek?
 (is that finished, the book? = have we finished that, the book?)
 (Sarah): nee, *hij* is nog niet afgelopen (Sarah, 2;5.9 week 128)
 (no he (=it) is not yet finished = we haven't finished it yet)

The naming function <+ref> of the demonstrative *dat* blocks its use as a presentational operator. The same blocking effect followed from the new rule that added an I⁰ with EPP effect to all predicates (section 3). The element *dat* turns from an utterance operator for presentationals into an obligatory marker for naming phrases (DPs). The merely pragmatic intention for applying a content sign as a name is now expressed by a functional category in syntax. The pragmatic <+ref> of previous diagrams can be spelled out as <+D>, see (21) above, and (24).

- (24) *Bootstrapping* D⁰
 Pragmatic intention of naming a topic
 ↓
 Topic optionally expressed by a gesture sustainable deictic operator F⁰
 ↓
 Obligatory presence of F⁰ equals D⁰ (reference)

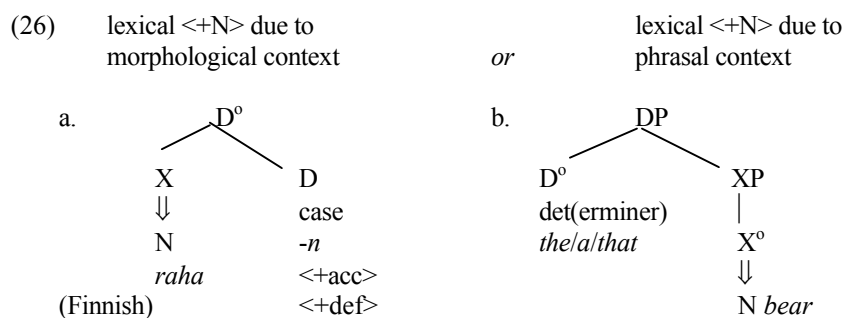
It is a remarkable fact that the acquisition graph for systematic D-marking coincides with the graph for free anaphors (25) shows. By free anaphors I mean (full/weak/clitic) pronouns as opposed to bound anaphors (reflexives).

- (25) Dutch Sarah: Acquisition of D-marking and free anaphors



Clearly <+D> stands for the referential function. The <+D> of personal pronouns (23)c takes the same argument functions as a fully lexicalized DP, see (22)c. As soon as the referential function is expressed by the functional category <+D>, it becomes possible to use <+D> independently as a free anaphor in argument position (Van Kampen 2003). Argument structure itself does not really exist without <+D/case> marking. Before D-marking, arguments are more like parts of a compound. They lack separate referentiality. Argument structure, say Baker's UTAH (Baker 1988:46f), needs <+D/case> marking. The UTAH (Uniform Theta Assignment Hypothesis) holds that theta roles select a linearly fixed X-bar configuration for the theta-assigning verb and its theta-carrying argument. Such a close relation between argument-hood and referentiality has been argued for by Williams (1994: chapter 6). See Evers & Van Kampen (2001) and Tracy (2002) for the role of the UTAH in establishing the argument position.

As before with the syntactic operator <+I>, the syntactic operator <+D> selects a subset of the category-neutral X^0 . The lexical category <+N> can now be deduced for all elements that appear as the complement of a determiner D^0 , see (26).



The context-sensitive acquisition rules in (26) demonstrate again how language-specific paradigms lead the learner towards universal properties. Note that my point of view deliberately contradicts the more common position that universal categories would be given and that they allow the identification of the more language-specific arrangements (Chomsky 1995).

5. Conclusion

There is a proto-grammar that still lacks functional structures and syntactic categories. Its structures are situation bound. Proto-grammar starts with pragmatic operators and with content signs X^0 that are category-neutral. They can be used, and are used in child language, as proper names with topic intention and as brand names with characterizing intention. Deictic operators for topic and deictic operators for comment are regularized as D-marking and I-marking. As soon as D-marking and I-marking have been perceived in the input, the topic-comment of proto-grammar turns into the subject-predicate of real grammar. These functional categories I^0 and D^0 are the crucial bootstraps for category assignment <+V> and <+N>, as a subdivision of the lexicon. The acquisition of language specific devices of I^0 and D^0 are the condition for distinguishing the universal categories V and N.

References

- Baker, M.C. (1988) *Incorporation. A Theory of Grammatical Function Changing*, Chicago: University of Chicago Press.
- Bellugi, U. (1967) *The Acquisition of the System of Negation in Children's Speech*, PhD. dissertation Harvard University.
- Blom, E. and E. Krikhaar (2002) 'Where to use a nonfinite clause? An experiment on the use of verb forms in early Dutch and English child language', in: J.Costa and J.M. Freitas (eds.) *Proceedings of GALA 2001* Palmela, Portugal
- Braine, M.D.S. (1963) 'The ontogeny of English phrase structure', in: *Language* 39, 1-13.
- Braine, M.D.S. (1976) 'Children's first word combinations', *Monographs of the Society for Research in Child Development* 41 (Serial no. 164).
- Chomsky, N. (1995) *The Minimalist Program*, Cambridge Mass./London: MIT Press.
- 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.
- Grimshaw, J. (1991) 'Extended Projection', ms. Linguistics and Cognitive Science Program and Center for Complex Systems, Brandeis University.
- Hoekstra, T. & N. Hyams (1998) 'Aspects of root infinitives', *Lingua* 106, 81-112.
- Jordens, P. (2002) 'Finiteness in early child Dutch', *Linguistics* 40 (4), 687-765.
- Kampen, J. van (1997) *First Steps in Wh-movement*, Delft: Eburon.
- Kampen, J. van (2001) 'Bootstraps at two for lexicon and discourse', in: Proceedings of ELA, Lyon, December 5-8, <http://www.let.uu.nl/~Jacqueline.vanKampen/personal/>
- Kampen, J. van (2003) 'The acquisition of the standard EPP in Dutch and French', ms.Utrecht University <http://www.let.uu.nl/~Jacqueline.vanKampen/personal/downloadables.htm>
- Kampen, J. van (to appear) 'Learnability order in the French pronominal system', to appear in: R. Bok-Bennema & B. Hollebrandse (eds.) *Selected Papers from Going Romance 2002*. Amsterdam Studies in the Theory and History of Linguistic Science.
- Lebeaux, D. (1988) *Language Acquisition and the Form of Grammar*, PhD. dissertation University of Massachusetts.
- Lyons, J. (1979) 'Deixis and anaphora', in: T. Myers (ed.) *The Development of Conversation and Discourse*, Edinburgh: University Press, 88-103.
- MacWhinney, B (1991) *The CHILDES Project: Tools for analyzing talk*, Hillsdale New York: Lawrence Erlbaum.
- Marantz, A. (1997) 'No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon', in A. Dimitriadis, L. Siegel, et al. (eds.) *University of Pennsylvania Working Papers in Linguistics* 4.2, 201-225.
- Pinker, S. (1984) *Language Learnability and Language Development*, Cambridge Mass/London: Harvard University Press.
- Powers, S. (1996) 'MAPping phrase markers', in: Ch. Koster and F. Wijnen *Proceedings of GALA* (Groningen Assembly on Language Acquisition), 303-313.
- Powers, S. and Lebeaux, D. (1998) 'More data on DP acquisition', in: N.D. Dittmar & Z. Penner (eds.) *Issues in the Theory of Language Acquisition*, Bern: Peter Lang, 37-77.
- Radford, A. (1990) *Syntactic Theory and the Acquisition of English Syntax: the Nature of Early Child Grammars of English*, Oxford: Basil Blackwell.
- 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. (1992) 'From the initial state to V2: Acquisition principles in action', in: J. Meisel (ed.) *The Acquisition of Verb Placement*. Dordrecht: Kluwer Academic Publishers, 333-370.
- Tracy, R. (2002) 'Growing (clausal) roots: all children start out (and may remain) multilingual', *Linguistics* 40 (4), 653-686.
- Williams, E. (1994) *Thematic Structure in Syntax*, MIT Press, Cambridge Mass.
- Wittgenstein, L. (1953/1975) *Philosophische Untersuchungen*, Frankfurt am Main: Suhrkamp, 3rd edition.