E-language, I-language and the Order of Parameter Setting*

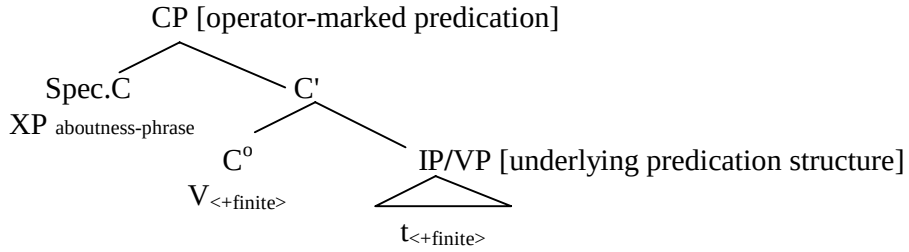
Arnold Evers and Jacqueline van Kampen

1 The problem: learnability of underlying structure

All the various grammatical characteristics of adult grammar will eventually appear in child language. They will do so successively, and reveal a learnability hierarchy that holds for all primary learners of the language. As a matter of fact Brown (1973:313ff) showed how 13 grammatical morphemes of English obey such an order of appearance. What we value in Brown is his attention to the order of learning step and his attempt to derive that order from the interaction of system factors and input factors. The first grammatical characteristics in child language are not arbitrary choices. They remind one of a Greenberg-like language typology. An explanatory relation between language typology, child language and universal grammar was brought to the fore in Jakobson (1942) for phonological features and, as he pointed out, the same relationship is likely to appear in syntax as well. Chomsky (1981) has proposed that primary language acquisition might be seen as a stepwise procedure that sets yes/no parameters in an a priori frame. This tallies well with Jakobson's conception of language acquisition. The parameters of grammar are set according to a kind of learnability hierarchy. The setting of some typological parameters should be understood as a precondition for the setting of other more language specific parameters.

Since child language acquisition is responsive to the typological features of the target language, it must be successful in picking up evidence from the input. However, there is a paradox here. The right selection of evidence must often abstract away from other construction specific and language specific features. For example, a so-called V-second language like Dutch moves all finite verbs into an operator position <+C>, where <+C> is defined as the first position c-commanding the predication IP/VP on the righthand side. The X^o <+C> head marks a prospective operator head and Spec,C a prospective operator phrase, as in (1). The V-second rearrangement of the finite verb V<+fin/+C> takes place only in root sentences and not in subordinates. Yet, as we will show in detail, practically all sentences addressed to the child are root sentences and these invariably have the finite verb in the <+C> position.

(1) a.



The unambiguous presence of the finite verb constitutes a problem for the acquisition of the underlying verb argument order. The reconstruction of the underlying argument order as VO or OV requires that the Dutch child abstract away from the ever-present finite verb in the second position. Miraculously, this is precisely what appears to happen. The child initially disregards the C-projection systematically. During the first year the child acquires a lexicon of predicative heads. Thereby, s/he sticks to 'root infinitives' and OV structures. S/he first

* The empirical load and general perspective of the present paper improved after presentations and discussions with audiences in Utrecht (1998-2000) and San Sebastian (1999). Especially useful were comments made by Jacqueline Evers-Vermeul, Gaby Herman, and Bill Philip. Remaining misconceptions are due to the authors.

establishes all verbal elements in the predicate final position of O-V and Adv-V frames. This raises the following questions. First, how did the child know in the first place that OV frames offer the best orientation? Second, how did the child come to realize that Dutch is a V-second language and that s/he should leave out the verb in second position? This can be phrased in a slightly more abstract way. A learning procedure cannot arrive at the underlying structure if it has to abstract away from the effect of rules that are still unknown. The relevant rules are unknown, since by assumption the typological form of the target grammar is still unknown.

The paradox is simple but far more significant than is usually realized. It raises problems for a whole line of acquisition studies. For example, it has been argued (Roeper 1972, repeated in Roeper and Weissenborn 1990, Penner 1993, Powers and Lebeaux 1998) that the underlying and predicate final position of the verb in V-second Dutch/German is learnable due to some awareness by the child that the basic order is present in finite subordinate structures. This is a bold conjecture. Subordinates are rare in child language. Only 2% of the mother's input structures were finite subordinates in a count we made. Moreover, subordinates do not appear in the child's output until the V-second rule for root sentences has been fully acquired. Yet, these are minor points as compared to the major questions. Suppose the child has a UG inspired disposition for the distinction between root and subordinate clauses. That assumption by itself does not answer the question by which formal syntactic distinction s/he can apply that distinction. Nor does it answer the question how s/he comes to the conclusion that the subordinates present the underlying form for the roots. There are other proposals for setting the OV/VO parameter as well. Lebeaux (1988:26) is skeptical about the subordinate argument. He proposes that the position of a non-finite verb to the right of the negation element is a better indication for the shape of the underlying structure. An early positioning of the negation element in child language was observed by Deprez and Pierce (1994). Lightfoot (1991) adds the particle stranded in the predicate final position. It might trigger the child's decision that the finite verb has to be predicate final as well. Nespor et al. (1996) assume that children may set the directionality parameters of their language on prosodic information alone, even before syntactic categories have been figured out. This works fine for VO Italian and OV Turkish. Yet, as they observe, the Dutch child is in a more precarious position. S/he can deduce the right-headedness of the VP only by paying attention to the stress patterns in certain subordinate periphrastic phrases (V-Aux patterns) with crucial knowledge of the syntactic category status.

All these acquisition studies imply that the learner might somehow first acquire certain language specific properties, and subsequently derive a typological parameter**.** In a sense, this is an anti-Jakobson view on the order of acquisition. The learner is assumed to have acquired or identified the categories and language specific properties of <+/-C> <+/-root structure>, <+/-V>, <+/-Aux>, <+/-fin>, <+/-Neg>, <+/-particle>, <+/-periphrastic phrase> in spite of the fact that there has not yet been a decision about the language type.[1] Subsequently, there is a remarkable hint from UG such as "Do not move your particle ever, no matter how little sense the particle makes by itself without the verb" (for the Lightfoot child). Or "Assume that the finite verb originates to the right of the negation elements, even if your parents persist on having it to the right" (for the Lebeaux child). Or "Pay all your attention to the subordinates first if you are so clever to spot them. Rare though they may be, they are preciously informative" (for the Roeper child).

[1] The problem is inherent to the logic of acquisition. So the problem of relating highly language specific circumstances and very general principles does not change if one were to assume that language acquisition takes place in a pre-production period, before the child has made any babbling at all. Nor will the problem clear up if one were to assume that at a very deep level of analysis all language types are parallel. Such moves may be correct as such, but they do not solve the acquisition problem, seen as a problem of selecting eventual evidence.

A completely different kind of parameter setting has been considered in Gibson and Wexler (1994). One that represents a far more daring 'full competence' view. They assume that the learning procedure starts with an arbitrary typological parameter setting. This initial parameter setting is provisional and highly flexible. Subsequently the provisional setting may enable the child to get the key and acquire the more language specific peculiarities. Gibson and Wexler (1994) offer an extensive analysis of their proposal.

1.1 Typological parameter setting: Gibson and Wexler (1994)

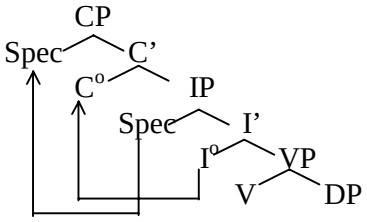
Gibson and Wexler (1994) fully realized the acquisition paradox, brought it into focus and proposed the following solution. Suppose all parameters are set provisionally, collectively and arbitrarily. Confronted with sentences from the target language, parsing failures will abound. Some of these parsing failures might be remedied by changing just one of the provisionally set parameters. Gibson and Wexler (1994) assume that this is the minimal step a language acquisition procedure can be prodded to make (Single Value Constraint), and that, if it does take this step, the parameter is set at a different value, again provisionally (Error-driven Change). Meandering around in the parameter possibility space, the language acquisition procedure might drop into a parameter setting realized in the target grammar. A parsing failure can no longer occur and in that sense the parameter setting has become irreversible. Gibson and Wexler subsequently raised the question whether such a full competence scenario will always lead to convergence or hardly ever. They came up with an artificial example and considered a language acquisition procedure that is, in its arbitrary ways, already considerably advanced. Six categories have been established {Subject, Object-1, Object-2, Adverb, Verb, Aux} and three either/or parameters are provisionally set (i) complement order (Object-V/V-Object), (ii) specifier order (Subject-VP/VP-Subject), and (iii) <+finite, +/- __C>, alias V-second (present or absent).

The three either/or parameters define eight Greenberg-like language types. By assumption, each of them could be the target configuration or the provisional source. The TARGET configuration is the typological parameter setting as realized by the language type that is to be learned and the PROVISIONAL SOURCE is the parameter setting as arbitrarily assumed by the language learner. If the two, source and target grammar, happen to coincide, some fortunate toddlers will find themselves immediately at home in their adult environment. Others by contrast will have to do a lot of resetting. Gibson and Wexler demonstrate that some types of the latter group (the real learners) cannot possibly meet with appropriate parsing failures such that these will always lead them on the right track. A trivial case is a source grammar generating a superset that includes the target strings, see (2). The V-second types in (2) and (4) below are represented by an underlying IP/VP structure, i.e. before the movement of <+fin>/Aux to C^o and the topic movement of some other constituent to Spec, CP.

(2) *Superset grammar* → *Subset grammar*
Source grammar: (**C^0**) Subject [Aux [Verb Object]] with V-second
Target grammar: Subject [Aux [Verb Object]] without V-second.

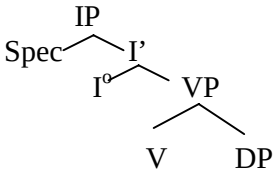
An, imaginary, example of a 'triggering' sentence in the target grammar of (2) is given in (3). It does not lead to any parameter resetting for the provisionally set source grammar.

(3) a. Provisionally set on (Source grammar)



b. Input sentence (English) '*You must eat your porridge*'. Analyzable by both the source and the target grammar.

c. Target grammar

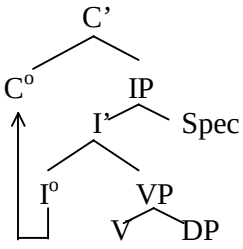


No parsing failure by the incorrect source grammar will occur. The V-second source grammar applied by the child handles the input quite well. It must prepose Aux to the <+C> position and may freely follow up by preposing the subject in a Spec,C. This faultlessly but incorrectly applies C-movement rules and simulates the SVO target grammar language without V-second. Additional data may help out. If the target grammar allows a topicalization movement 'Adverb-Subject-Verb' and the source grammar insists on 'Adverb-Verb-Subject', the target string set is no longer a subset of the source set of strings and the <+ V-second> setting can be reset to <– V-second>. The erroneous source grammar expects '*tomorrow [must [you eat your porridge]$_{IP}$]$_{CP}$*' (Swedish type) but the input corrects with '*tomorrow [you must eat your porridge]IP*' (English type). Gibson and Wexler add the Adverb preposing rule. Their intention is to keep the subset/superset problem out of their model. They also keep out the C-movements of <+wh>/<+Q> marked phrases. They have a more interesting problem of parameter setting in mind.

This more interesting problem appears for cases of non-intersecting distributions, like in (4). Again, an imaginary example is given in (5).

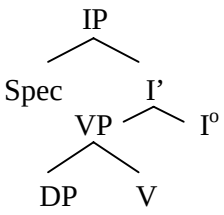
(4) *Non-intersecting grammars*
Source grammar: (**C°**) [Aux [Verb Object] Subject with V-second
Target grammar: Subject [[Object Verb] Aux] without V-second.

(5) a. Provisionally set on (Source grammar)



b. Input sentence '*You your porridge eat must*'. Not analyzable by the source grammar, but no resetting possible.

c. Target grammar



If the target grammar offers its input sentence (5)b, a string Subject-Object-Verb, the obvious conclusion must be that the V-second cannot be right. Unfortunately, this conclusion cannot reach the parameter resetting center, since it is not endorsed by solving a parsing failure. Dropping V-second from the source language yields no more than an Aux+Verb initial structure, where a verb final is needed. Inevitably suffering the unresettable but inane V-second, one cannot solve parsing failures by switching the complement order either, since the subject final position of the source language continues to be unacceptable for the target grammar. Nor will switching the subject position improve the complement order or the verbal distribution. Hence, no single parameter resetting will ever solve a parsing failure. A source configuration as the one in (5) above, where the learning scenario is bound to fail, is said to be a 'local maximum' with respect to the target grammar.

Gibson and Wexler found six such cases of local maxima and all involved a source grammar with V-second and a target grammar without. The (fictitious) scenario has the following morbid implication. Suppose local maxima may arise from this type of parameter setting and a young speaker toddling through the UG/parameter maze is lured into a positive setting of his V-second parameter, say due to a wh-question with subject inversion ('*what porridge must you eat?*'). Then s/he might quite well get trapped into one of the local maxima and suffer a remarkable, but so far unattested, type of typological language impairment. Gibson and Wexler argued that, obviously, setting the V-second parameter should not be considered until the directionality parameters for subject and object have been set in an irreversible way. This may help out, be it only within the demonstration space of six categories and three parameters. In principle, the entire formal exercise was set up by Gibson and Wexler to check the feasibility of a general language acquisition scenario. Subsequently, the existence of 'local maxima' was discovered and it turned out that the scenario may often allow but cannot guarantee the learnability of any established language type. Therefore, the basic result from the Gibson/Wexler exercise, as we see it, must be that the original learning scenario was shown to be not feasible, barring ad hoc measures. It is of course possible to add computational power. However, it must also be possible to add new paradoxical data sets.[2] It may be that neither move is useful. The relevant point may be neither to pull the Gibson and Wexler model out of its difficulties by brute power, nor to push it back in by new paradoxical problems. The relevant point may be elsewhere. It seems that merely grammatical hints in I-language will not clarify how the child, when confronted with the raw data, effectively solves the acquisition problem. Even an explicit full competence view unbelievably well-informed by a priori assumptions has been proved to fail. All the same, there is no doubt that almost any three-year old child succeeds in constructing a path from major typological decisions to language specific niceties.

[2] See for instance Berwick and Niyogi (1996), Frank and Kapur (1996), Fodor (1998, 2001).

It is repeatedly suggested by Gibson and Wexler that a parameter setting order might evade the local maxima and that such a convenient effect might follow from some marking hierarchy. They suggest that certain rules, for example the obnoxious V-second rule, are not an option for the acquisition procedure, at least not for a certain amount of time. Certain rules may require more computational space than the young speaker can muster. After all, you have to learn how to crawl before you learn how to walk. Developmental assumptions are quite plausible, as stressed again in Wexler (1998). That point is granted. The next point is: does it help? And if so, how? Suppose for example that all input is first processed by a non-transformational grammar, i.e. one without underlying structure. Then, there has to come a situation and a moment that will trigger the later movement solutions. The question now is, what will trigger the transformational reanalysis. Suppose, by contrast, that the initially impoverished acquisition procedure processes only a part of the input. Then no reanalysis of the transformational structures is needed. They have been left out. The question, then, is rather how the non-transformational part of the input has been selected. Gibson and Wexler do not enter into these questions. They mention the possibility of an initial but temporary (neural) handicap as a kind of blessing in disguise. Further, they keep the somewhat mysterious ordering to a minimum. It is here that we depart. What we suggest instead is that the ordering of parameter setting might be maximal.

1.2 The present approach: selection of evidence

Parameter ordering may result as an automatic effect from the system that is being confronted. According to the original form of this view by Jakobson (1942), the internal hierarchy of grammatical markings is bound to reveal itself as a learnability hierarchy. If one considers syntactic movement structures, it is clear that these are not simply structures with a different distribution. They are derived from an underlying downstairs structure. The movement structures are not acquired unless they have been perceived as having that derived structure. Hence, movement structures have an internal hierarchy of grammatical marking that leads from an underlying structure towards the derived one. If so, the primary input should preferably allow the perception and acquisition of the underlying structure before the movement complications are considered. The acquisition procedure needs an EVIDENCE FILTER to scan the primary input, such that direct evidence for underlying structure is separated from evidence for derived structure. An empirical approach to study the selection of evidence might run like in (6).

(6) a. Find clear cases for the acquisition of movement rules
b. Reconstruct the decision procedure for the underlying structure

Mere descriptions of incremental steps have been made earlier (Lebeaux 1988, Radford 1990, 1996, Clahsen 1991, Wijnen 1995). What we have in mind here is a decision procedure that derives the linear order of the learning steps as in Van Kampen (1997). The present paper will argue for this approach by an acquisition analysis of the C-movement rules, move <+wh/+C> (wh-phrase movement) and move <+fin/+C> ((residual) V-second head movement) in Dutch and English. Both rules in both languages apply to the predication IP/VP. In fact they extract material from the IP/VP. It will be demonstrated how the learning procedure reconstructs the underlying structure of these two movement rules in Dutch and English.

Section 2 compares the C-projection in Dutch and English root clauses. The English variant of V-second is also known as ‘residual’ V-second. Acquisitional graphs will show later on (sections 3 and 4) that English residual V-second is much harder to learn than its Dutch counterpart. The English head movement requires almost 60 weeks and the Dutch one almost 20 weeks. Section 3 for Dutch, and section 4 for English make clear how the same

natural evidence filters are at work in the acquisition procedure of both languages. Due to the different input, the evidence filters select almost immediately a different basic order. Subsequently, they provoke real but wildly different learning paths to more or less parallel <+wh> C-projections. The conclusion in section 5 states the following three points.

(i) *Facts*
There is a parameter setting order. The order and the relative speed of parameter setting can be demonstrated by linear acquisition graphs of the type introduced by Brown (1973). The actual order of parameter setting is as expected by Gibson and Wexler (1994). Paradoxically, underlying directionality comes first.

(ii) *Acquisition procedure*
The acquisition order and the acquisition speed of move <+wh/C> and move <+fin/C> in Dutch and English are different but due to the same evidence filters on the input. These are quantitative filters that depend on frequencies in E-language. They allow an uninformed language acquisition procedure to perceive underlying constructs first and their transformationally derived variants thereafter. The incremental stages follow again from I-language hierarchies highlighted by E-language frequencies.

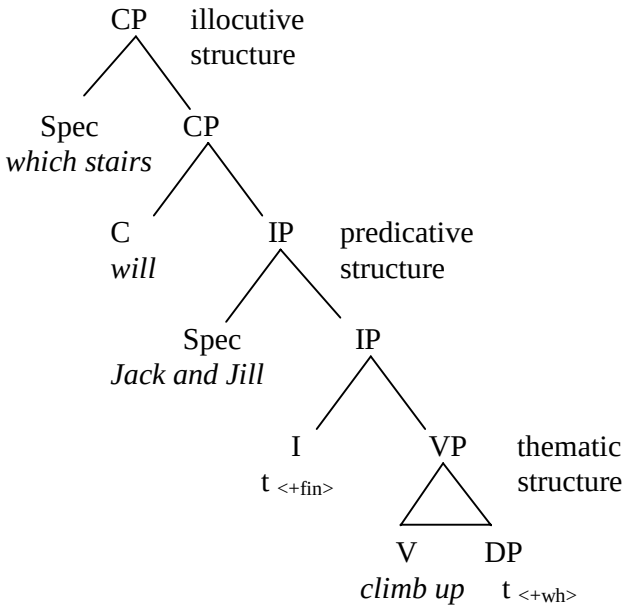
(iii) *Metaphysics*
The learnability of natural grammar follows from an interaction between I-language and E-language. Both factors are imposed on the learner. The present proposal stresses control by input and shies away from Chomskyan mentalism where wired-in properties of the human brain itself lay down their a priori conceptions on a cultural product.

2 Parameter hierarchy and root clauses in English and Dutch

This section starts the program indicated in (6) above. First, a parallel will be sketched between the C-projection in English and Dutch with attention to the more restricted nature of the C-projection in English. Thereafter, the major acquisition differences will be indicated that exist between the rise of the English C-projection as compared to the Dutch one. The relevance of this difference for a theory of language acquisition is shortly indicated.

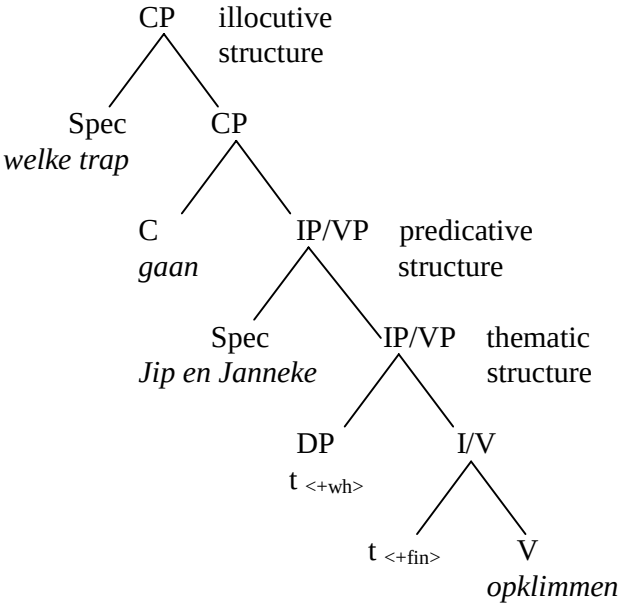
The acquisition of move <+wh> and move <+fin> in Dutch and English is relevant to (6)a and (6)b. Both languages extract the wh-constituent out of the IP/VP and move it into the initial Spec,C position and both languages bring the finite form <+fin> into the C-head position, cf. the diagram in (7) for English.

(7)



The Dutch structure is parallel to the English one, but there are four differences. Firstly, the Dutch C^o-position is not restricted to a set of auxiliaries as it is in English. Any finite verb in Dutch will enter the C^o-position in root clauses, although the token frequency of auxiliaries runs high. In section 3.2, an analysis of the input will show that they constitute almost 70% of the finite verbs. Secondly, the Dutch Spec,C position is not restricted to <+wh> constituents as it is in English. Any major constituent in Dutch may enter Spec,C in root clauses, provided it does not contain the finite verb and is not <–definite>. Thirdly, all root sentences in Dutch are to be marked by the finite verb in the position in C^o and an 'aboutness' phrase <+wh> or <–wh> in Spec,C. A fourth difference is that the <–fin> verb in Dutch occupies a predicate-final position, see diagram (8).

(8)



The acquisition of the structures (7) and (8) raises two issues of parameter hierarchy.

The first issue is that the acquisition of the C-projection cannot be a mere parroting of PF distributions. The learning procedure must detect that the elements in the C-projection

relate to empty places ($t_{<+fin>}$ and $t_{<+wh>}$) within the IP/VP phrase. For that reason, the C-projection cannot be acquired before the major IP/VP characteristics have been established. This is a parameter ordering. It is expressed in (9) (see also Van Kampen 1997:10ff,160ff).

(9) UG: VP parameters $<_{precede}$ IP parameters $<_{precede}$ CP parameters

The Dutch root sentence offers no position identifiable as I^o. The only lawful inhabitant of the I^o position is the <+fin> head. Precisely the <+fin> head is moved up to C^o in root sentences as illustrated in diagram (8). Since the primary linguistic data as present in root sentences offer the learner no evidence for an I^o position, they offer no evidence for directionalities in the I^o position either, see Watanabe (1993:15, 1994). This spells trouble for an input controlled parameter setting that restricts itself to root questions and nevertheless wants to solve the typological orientation first. Roeper (1972) and all the people in his wake clearly realized this point. We will stick to input control by root sentences only. Section 3 will turn the problem into a stronghold of our approach by solving it. In the mean time, we beat around the bush by making no distinction between the V-projection line and the I-projection line in root sentences.[3] The diagram in (8) interprets the initial constituent root clauses as the topic, and the initial structure as a C-projection added to a predication IP, maybe VP.[4]

The second issue is the acquisition of the Spec,C position. When the head C^o is projected, the preposed wh-element can be identified as a Spec,C. This suggests the acquisition order $C^o <_{precedes}$ Spec,C. However, a general non language-specific order is not supported by the facts. As we will see below (sections 3 and 4) the general picture for Dutch in English acquisition of CP is as schematically represented in (10) and (11).[5]

(10) Dutch: move <+fin> to C^o $<_{precedes}$ move <+wh> to Spec,C

(11) English: move <+wh> to Spec,C $<_{precedes}$ move <+fin> to C^o

It is worthwhile to pay attention to the difference in (10) and (11). Grammatical characteristics are acquired in a certain order, which perhaps may be read as 'parameters are set in a predetermined linear order'. The parameter setting order can be language specific in principle. Hence, many more statements like (10) and (11) will be needed to describe the language specific path of the acquisition procedure. The question is what forces the order of the acquisition steps. One would not prefer to postulate a new network of (biological) a priori constraints in acquisition theory. If the acquisition order is language specific in principle, one might rather consider the following way of looking at things. Children that want to climb a

[3] There is a variety of positions, that we cannot do justice to here. Weerman (1989) argues for the absence of a separate I-projection line in adult Dutch.

[4] For optional functional projections see also Grimshaw (1993). Most acquisition studies insist on a universally given IP. See e.g. Clahsen (1991:384) for an attempt to reconstruct an IP in German by means of the subordinate structure. Other solutions fit into the antisymmetric syntax perspective (Kayne 1994, Zwart 1996). Weissenborn (1991), for instance, has suggested that the initial structure is a left-headed IP: Subj I^o VP (Weissenborn 1991), because child language has a relative preference for the subject in sentence initial position. At the same time, non-subject topics are not excluded and appear as early as the rise of V-second in C^o (week 107 for both Sarah and Laura, CHILDES, Van Kampen corpus, see section 3.3). Poeppel and Wexler note that the child Andreas (age 2;1) had already 30% non-subject topics in finite sentences. Something may play a role in the child's initial preference for a subject topic. Non-subject topics are a discourse phenomenon. The neutral and isolated sentence is more likely to have the subject as its topic. Hence, this is a more likely form for the beginning speaker as well.

[5] The late acquisition of wh-movement in V-second languages has been noticed before. See for empirical evidence that V-second is acquired before wh-movement in child Dutch: Van Kampen (1992, 1997); in child German: Tracy (1994); in child Swiss Bernese: Penner (1994); in child Swedish: Santelmann (1995).

flight of stairs, - also a challenge of that period -, must start with the first stair, no matter whether they are or are not a priori informed about the nature of stair-climbing. The order of learning activities is imposed by the structure of the thing that is being confronted. There need be no more to it. The basic simplicity of a device and the ease with which it is mastered need not in any specific way reveal the neuro-motoric or cognitive complexity the child is blessed with. If the child can make it to the first step of the stairs, there is almost immediately the sensational accomplishment of getting to the next floor. That sensation owes something to the effectiveness of the staircase as an autonomous device. Neural organization as such does not tell us about properties of the "staircase effectiveness", nor will staircase principles inform us very much about neuro-motoric organization. These are interacting systems, not mutual reflections. The staircase logic may hold for language competence and language acquisition as well. The human neural system, adaptable and rule-hungry as it is, succeeds in tracking down the set-up and the rules of some external system, not unlike what may happen with board-games, field-games, square-dancing, or whatever other 'artificial' cultural constructs. Subsequently, it simulates, so to speak, that external system in its own activities. The limited number and parallel nature of grammatical options, as aimed at in the principles/parameters system, need not follow from any specific biological constraint. It may follow as well from some elementary comparative design effectiveness. Inventions fit into a hierarchy of their own. The decimal notation of fractions was invented by the Dutch mathematician Simon Stevin in his arithmetic book *De Tiendhe* (1596, 'The Tenth One') dedicated to all merchants and craftsmen. In fact, Stevin had reinvented the decimal notation for fractions. It appears as a calculation trick in the work of Islamic mathematicians before 1000. The famous poet and mathematician Omar Khyam (around 1100) raised the question whether ratios were not a family of numbers rather than an operation. The decimal notation of fraction is explicitly introduced and used in Al-Kashi's work *Sallam At-sama* (1407, 'The Stairway of Heaven'). Al-Kashi was the leading mathematician and astronomer at the Mongol court in Samarkand.[6] By the logic of the invention, one cannot hit upon the decimal notation of fractions before there is a decimal notation and arithmetic of natural numbers. The Islamic astronomers, the Dutch engineer and in fact every later high school student had to follow the same order of learning steps. It would be strange if a hierarchy of discovery held for rules and categories in arithmetic, but not in grammar.

The view we take here is that the inventions of the mind, - grammar for example -, are to be understood comparatively and in their own terms, rather than by invoking neural or behavioral constraints. This is meant to reaffirm De Saussure's (1915/1964:25) point of view ("il faut se placer de prime abord sur le terrain de la langue et la prendre pour norme de toutes les autres manifestations du langage"). The distinctions and solutions in the grammar are to be understood on their own terms. They are embodied in, but cannot be derived from or explained by the neural or social environment they function in. This must be close to the "bare conceptual necessity" or "perfect nature" of combinatorial systems spoken of by Chomsky (2000). Yet, Chomsky often implies something different than mere autonomy of grammar. He characterizes grammar itself as 'biological', because it is embodied in a human (species specific) neural system and he suggests that grammar, rather than being learned, 'grows' into the mind. We certainly do not advance this perspective. For example in ballroom dancing, the tango requires far more exercise than the Boston two-steps. This follows from its choreographic design, and does not merit the specific attention of a neurologist. In the same vein, getting the root question in English requires almost three times as much exercise as

[6] This information is taken from the website of the School of Mathematics and Statistics, University of St. Andrews, Scotland. URL: http://www-history.mcs.st-andrews.ac.uk/history/References/Al-Kashi.html

getting the Dutch root question. This fact as such will be argued below, in sections 3 and 4. Again, the difference may follow from design, grammatical design, and does not merit the specific attention of neurologists.

There is nothing strange in considering universal grammar as a hidden hand behind the learnability of grammar without being committed to the position of a biological grammar. This answers the worries above that statements like (9), (10) and (11) may imply new a priori biological constraints. They need not. Rather, one may consider language acquisition as input controlled. One may hold, as we will do, that incremental learning is crucially dependent on quantitative proportions of the input, without being committed to the position that children are merely parroting high frequency distributions. In fact, they are not. The highly frequent <+fin> marking, more than 90% of the input (adult Dutch), is largely disregarded by the child. Less than 10% in the output (child Dutch) mimics a V-second phenomenon. This is more or less stable until the argument structure that V-second applies to has been acquired. We will now spell out that this factual claim in quantitative detail and explain it.

3 Evidence filters and the Dutch order of parameter setting

The first section (3.1 'Argument directionality') demonstrates how the underlying argument order for Dutch (Subject before VP; Complement before V) can be derived from the mother's input. The mother's input structures are at face value remarkably non-informative (table (12)). That picture can be changed by applying Lebeaux' (1988) theta/Case criterion as an input filter (table (23)). The second section (3.2 Apparent V-second) points out how the transformationally derived C-structures stay out of the child's preliminary non-transformational grammar. Some V-second structures seem to make a marginal appearance in the earliest output data. These V-second structures need not be interpreted as premature movement structures. They are better understood as alternative theta/Case frames. We realize quite well that full competence hypotheses claim a very early awareness by the child for all functional categories, not only <+fin>/V-second>, whereas we deny it. The present research program is a rival to the full competence program. The third section (3.3 'Longitudinal graphs') presents the longitudinal analysis for the acquisition of the transformationally derived V-second. It points out four significant acquisition facts, facts that need to be explained. The fourth section (3.4 'Reconstruction of the verbal chain') will explain the four facts from the preceding section and in principle solve the Gibson and Wexler paradox. It is argued that the preliminary non-transformational grammar gives rise to empty places in the analysis of transformational input. This leads to an input driven construal of A-bar chains, the rise of adult grammar. The input driven discovery procedure is throughout dependent on two factors. Firstly, there is a preliminary grammar from earlier learning steps (the I-language factor), which is reaffirmed by 70% of the data. Secondly, there are various percentages of new data (the E-language factor) that fit the preliminary grammar but for the movement rule. Hence, the analysis requires an attention to the order of learning steps (the development of the I-language factor) and input percentages (the pressure of the E-language factor). A final section (3.5 'V-second $<_{precedes}$ wh-movement') demonstrates by means of longitudinal graphs, that move <+wh> follows move <+fin> in the acquisition of Dutch. A fact that will acquire significance in section 4 where it will be compared to the acquisition of C-movement rules in English.

3.1 Argument directionality: the raw input data

Gibson and Wexler (1994) strongly suggest that the acquisition procedure should somehow not set the V-second parameter before it has fixed the directionality parameters. Fortunately, the Dutch child does indeed set the complement parameter at 'Complement before V', and the

specifier parameter at 'Subject before VP', before addressing the V-second parameter.[7] How did the Dutch child know about the dangers of the Gibson/Wexler paradox? Considering the quantitative relations within the input, the primary choices are remarkable. A simple analysis of the primary input from the mother in the Van Kampen corpus in CHILDES (MacWhinney 1991) shows that the child's acquisition procedure is highly selective in its approach to the available evidence. We assume that the mother's input in the files will roughly reflect the proportions of the input in general. The input given by the mother was taken from three consecutive files just before the child acquired the V-second rule. The mother's input figures for the complement order, specifier order and V-second, as analyzed in (12) to (17) below, have been taken from the recordings of conversations between Sarah and her mother. The mother's input in these three files contain 1017 structures with a verb and these offer simultaneous evidence for all major properties of Dutch grammar. The question is how the evidence can be selected by the child's acquisition procedure. We concentrate for the moment on evidence for the three Gibson/Wexler parameters, (i) complement order, (ii) specifier order, and (iii) +/- V-second.[8]

Following Van Kampen (1997:chapt.2), we categorize the three Gibson/Wexler parameters as VP-argumental, IP-predicational and CP-illocutional. The distributional evidence in (12), then, seems available to the acquisition procedure. Each cell in column A indicates the percentage of the mother's utterances that provides positive evidence for the setting of the three parameters. Each cell in column B indicates the percentage of the mother's utterances that provides contradictive evidence for the setting of the parameter. Assuming that the child pays as much attention to this positive evidence that a given parameter is not set in a certain way (column B) as to that suggesting that it is (column A), the Dutch child should conclude that: there is V-second, the subject follows the VP and the object follows the verb. In column C we list the percentages of the ratios A/A+B and B/A+B.

(12) The Gibson and Wexler parameters: preliminary input percentages
Out of 1017 utterances with a verb (mother Sarah, Van Kampen corpus)

Parameter	A [+ evidence]	B [- evidence]	C ratios A/A+B	B/A+B
a. *CP-illocutional* V-second	93%	7%	93% (93/100)	7% (7/100)
b. *IP-predicational* Subject-VP	20%	49%	29% (20/69)	71% (49/69)
c. *VP-argumental* Object-V	15%	18%	45% (15/33)	55% (18/33)

The percentages of columns A+B in (12)b and (12)c do not reach 100%, because not all sentences have a subject or an object. To compare the relative weights of +/- evidence for Subject-VP and Object-V order, we have added column C.

[7] For instance, a check of the Sarah corpus showed hardly any deviation by Sarah from primary directionality parameters in the first ten files before the rise of V-second (week 81-110, Van Kampen corpus, CHILDES). See also Clahsen (1982) for German, Schlichting (1996) for Dutch.

[8] These parameters presuppose an identification of syntactic categories, as well as the notions subject and predicate. This preliminary issue is not addressed here. See Van Kampen (1997, 2001) for further explorations of this matter.

V-second

As seen in (12)a, column A, no less than 93% (915n) of the mother's utterances offer direct PF evidence for a Vfin in first or second position. In 66% of the cases the Vfin was a modal/auxiliary, see (13)a, and in 27% of the cases it was a lexical verb, see (13)b.

(13)	a. (XP) V<+fin,+aux>		66%
	papa moet een boek lezen	(daddy must a book read)	
	b. (XP) V<+fin,–aux>		27%
	papa leest een boek	(daddy reads a book)	

Among the 7% without a V-second structure ((12)a, column B), there were 2% subordinate structures with a finite verb in non-V-second position and 5% elliptic non-finite structures.

Subject-VP

The relevant number of structures for the subject evidence, see (12)b, drops to 69% of the mother's utterances carrying a verb, i.e. 697 of a total of 1017 utterances. Some 31% is considered irrelevant, mainly because there is no (overt) subject.[9] [10]

The subject inversion [(XP_{topic})-$V_{<+fin>}$-Subject] has complicated the subject distribution in 49% of the 69% set, see (12)b, column B and the examples in (14). The non-subject topic XP may be missing (in yes/no questions and topic-drop structures). The Vfin in (14) represents the items V<+fin,+aux> and V<+fin,-aux>.

(14)	[(XP)-Vfin-Subject-(rest)]		49%
	(nu) moet papa een boek lezen	((now) must daddy read a book)	
	(nu) lees papa een boek	((now) reads daddy a book)	
	leest papa een boek?	(reads daddy a book?)	

Only 20% of the 69% set offers straight PF parameter evidence for a *Subject-VP* adjacency, where the VP includes the <+fin> marking, see (12)b, column A. This 20% consisted of 18% root sentences ((15)a), and of 2% subordinate structures ((15)b).

[9] This 31% consisted of the following structures.

(i)	imperatives	12%	
(ii)	elliptics and subject-drop	7%	
(iii)	stereotype naming phrases	12%	(types: *wat is dat?/ dat is X* 'what is that?'/ 'that is X')

[10] It has been objected a few times that the child might insert empty objects and subjects. One may calculate that this by itself will not remove the input paradoxes, but the objection misses a more important point. The present approach is highly stingy about the introduction of empty place. Empty places cannot be introduced by the present approach before the directionality parameters have been firmly established. These cannot have been set a priori, because they are the very basis of typological variation. In short, there are no empty places in the beginning system.

(15) a. [Subject + [Vfin-(rest)]$_{Predicate}$] 18%
papa moet een boek lezen (daddy must a book read)
papa leest een boek (daddy reads a book)
b. [Subject + [(rest)-Vfin]$_{Predicate}$] 2%
ik wil dat papa een boek gaat lezen (I want that daddy a book goes read)
ik wil dat papa een boek leest (I want that daddy a book read)

Object-V

The relevant structures for setting the Object-V parameter drops to 33% of the mother's utterances carrying a verb, 334 of a total of 1017 utterances. See (12)c. This is because the Object-V parameter can only be set by transitive structures, verbs accompanied by their complement (direct or prepositional). The other 67% of verb-containing utterances is considered irrelevant. It consisted of intransitive structures or transitive structures with object (topic) drop.

The straight Object-V predicate pattern was muddled somewhat by sentences of the three types shown in (16) in 18% of the 33% set, see (12)c column B.

(16) a. [Object-V$_{fin}$-Subject] (topicalization of the object) 8%
een boek leest papa (a book reads daddy)
b. [(XP)-Vlex$_{fin}$-Object] (movement of Vlex to first/second posit.) 9%
papa leest een boek (daddy reads a book)
c. [(XP)-Aux$_{fin}$-Object] (bare auxiliaries) 1%
ik kan dat (I can that)

Only 15% of the 33% set offer straight PF parameter evidence for *Object-V* adjacency order, see (12)c column A and the example in (17).

(17) a. [(XP)-Aux$_{fin}$-Object-V] 15%
papa moet een boek lezen (daddy must a book read)

If acquiring a grammar were primarily something like getting the habit for a set of syntactic templates, each containing a fixed order of categories, the input characteristic in (12) are paradoxical indeed. The V-second evidence is overwhelming ((12)a, column C), but, as we will see, it is not even considered by the child before the VP parameters have been set. The Object-V parameter is set almost immediately, though the evidence for it trails that of other transitive structure distributions by 10% ((12)c, column C). Finally, how has the Subject-VP directionality been set? Part of the 49% of counter-evidence in subject-inversion structures, ((12)b, column B), has the subject in the specifier VP position with Aux-in-C^{o}: [(XP)-Aux-Subject-Predicate]. This would identify the subject position if the parsing mechanism of the acquisition procedure could abstract away from the C-movement rules. Obviously, this is too awkward an assumption. A learning procedure should find out about the C-movement rules, not presuppose them. In fact, this was the problem confronted by Gibson and Wexler. They showed how movement rules confuse the evidence about the underlying directionality and even their fairy-tale escape by means of arbitrary parameter setting and stumbling into the right solution could not guarantee the desired results. Yet, these results, {Subject-VP, Object-V, +V-second}, are unfailingly arrived at by Dutch three-years-olds. Hence, there must be another way out.

3.1.1 Evidence filters

As observed by Lebeaux (1988:11f), the child's initial two-word structures represent minimal theta-assigning government structures. This insight has been reapplied in Van Kampen's (1997) proposal that the language acquisition procedure starts by filtering the PF input for structures that directly reflect LF predication and theta relations, see (18).

(18) *predication relations*

a. beertje slapen	(teddy-bear (must) sleep)
b. deur open	((the) door (is) open)
c. papa lief	(daddy (is) nice)
d. boot daar	((the) boat (is) there)

theta relations

a. boekje lezen	((I want to) read (a) booklet)
b. banaan eten	((I am going to) eat (a) banana)

If the learning procedure is attentive to rigid government structures for theta assignment and predication, content elements are perceived as belonging to linearly fixed frames for theta assignment. The learnability of content elements is probably served thereby (Gleitman 1990, Brent 1994). Brent (1994) added a computational simulation of this idea. The choice of the frame follows from the most frequent PF predicate form. There is a double acquisition advantage once a uniform theta frame has been established.

Firstly, semantic oppositions between the lexical content phrases can more easily be perceived as standard distinctions (affected object, means, location, effected object) added to a fixed syntactic frame. This point tallies well with Hale and Keyser (1992), who argue that potential theta distinctions are standardized by the austerity of X-bar frames.

(19)

X	V	
a. boekje	lezen	(booklet read)
b. bal	gooien	(ball throw)
c. paardje	rijden	(horsie drive)
d. sap	drinken	(juice drink)

Secondly, PF variations in the linear order of the frame, as in (20), attract the attention and can be given additional uniform syntactic interpretation.[11]

(20)

$V_{<+fin>}$	X	
a. leest	boek	($read_{<+fin>}$book)
b. gooi	bal	($throw_{<+fin>}$ ball)
c. rijdt	paardje	($drive_{<+fin>}$horsie)
d. drink	sap	($drink_{<+fin>}$juice)

The Lebeaux frame explains Baker's Uniform Theta Assignment Hypothesis (UTAH) (Baker 1988:46f). Baker's UTAH holds that theta roles select a linearly fixed X-bar configuration between the theta-assigning verb and the theta-carrying argument.[12]

[11] The frames in (19) and (20) refer to the category V. This is to simplify the exposition. As argued in Van Kampen (1997: chap.2,3) the category system is not accessible that early. The contention there, as well as in Van Kampen (2001), is that all content signs fall under a category-neutral X. A step X → V cannot be taken before the main oppositions of the verbal paradigm are clear, i.e. not before the acquisition of move <+fin>.

The primary directionality relations reflect the underlying structure and they follow from a way of learning lexical content phrases and setting up the lexicon. Distributions of transitive structures, 33% of the predicates, are given in (21). The total of 33% is from (12)c.

(21) Object-V/V-Object parameter: input percentages on evidence filters

Evidence filter for Complement-V	A Object first [+ evidence]	B Object final [- evidence]	C ratios A/A+B B/A+B
a. + *Adjacent theta-assignor*	O-V 15%+3%	V-O 5%	**78%** (18/23) **22%** (5/23)
b. – *Adjacent theta-assignor*	(O-Aux-S-V 3%)	(V-S-O 4%)	
c. *No theta-assignor*	(O-Aux-S 2%)	(S-Aux-O 1%)	

a. ik ga **een boek lezen** (15%) — ik **lees een boek** (5%)
(I am going a book read) — (I read a book)
een boek lees ik (3%)
(a book read I)

b. een boek *ga ik* lezen (3%) — (dan) lees *ik* een boek / lees *jij* een boek? (4%)
(a book I am going to read) — ((then) read I a book) / (read you a book?)

c. dat kan ik (2%) — ik kan dat (1%)
(that can I) — (I can that)

If the language acquisition procedure is initially attentive only to the narrow frame (21)a shaped by adjacency plus theta assignment, the Object-V (18%) outnumbers the V-Object (5%). That amounts to 78% over 22% within this narrow frame, as given in (21), column C. The remaining transitive structures (21)b and (21)c would be irrelevant. In (21)b the OV or VO relation is not string adjacent. The subject is in between. Alternatively, in (21)c, there is no clear object theta relation. It is not claimed that the cognitive content of the theta role is hereby acquired. Rather there is a pragmatic cognitive content that now is being standardized due to a fixed syntactic position (abstract Case).

A parallel evidence filter can be used to determine subject directionality, where the predicate functions as the theta assignor, see (22). Subject directionality holds for verbal and non-verbal predicates (copular constructions). It is unlikely that early child grammar can distinguish these two kinds of predicates before the verbal paradigm is available (see Van Kampen 1997, 2001). Therefore early child language automatically addresses the more abstract Subject-Predicate directionality rather than Subject-VP directionality. The total amount of predicates in (22), 69%, is from (12)b.

[12] For the sake of clarity, it is not Baker's position that UTAH is inferred from or provoked by the input. He rather sees UTAH as an a priori frame that is imposed on the input by the learner, such that polysynthetic Mohawk derives its general object incorporation from an underlying VP that is never PF realized.

(22) Subject-Predicate/Predicate-Subject parameter: input percentages on evidence filters

Evidence filter for Subject-Predicate	A Subject first [+evidence]	B Subject final [-evidence]	C ratios A/A+B B/A+B
a. + *Adjacent theta-assignor*	**34%**	**16%**	**68%** (34/50) **32%** (16/50)
+ verbal predicate	(Aux)-S-VP 24%	V-S-(O) 16%	
- verbal predicate	(Cop)-S-XP 10%	---	
b. - A*djacent theta-assignor*	**13%**	**1%**	
+ verbal predicate	S-Aux-VP 7%	VP-Aux-S 0%	
- verbal predicate	S-Cop-XP 6%	XP-Cop-S 1%	
a+b combined	**47%**	**17%**	**73.5%** (47/64) **27.5%** (17/64)
c. *No theta-assignor*	(S-Aux 0.5%)	(Aux-S 4%)	

a. *verbal predicate* (24%)
(nu) moet **jij** (**een boek**) **lezen**
(now) must you (a book) read
jij leest (**een boek**) (you read (a book))
non-verbal predicate (10%)
(nu) is **het boek uit/ de jongen stout**
((now) is the book out/ the boy naughty)
ik vind **jou lief** (I find you sweet)

verbal predicate (16%)
(nu) **lees jij** (**een boek**)
(now) read you (a book)
lees jij (**een boek**)? (read you (a book)?)

b. *verbal predicate* (7%)
ik ga (**een boek**) **lezen**
(I am going (a book) read)
non-verbal predicate (6%)
dat boek is **leuk**
(that book is funny)

verbal predicate (0%)
(**een boek**) **lezen** ga **ik**
((a book) read am going I)
non-verbal predicate (1%)
stout is **die jongen**
(naughty is that boy)

c. ik kan wel (0.5%)
(I can indeed)

(dat) kan ik wel (4%)
((that) I can indeed)

The narrow frame (22)a (column C) assumes that a preceding functional category Aux or Copula can be neglected. The Aux and Copula elements are a-typical in the child's early two-word production. Hence, we feel that the child is unaware or insufficiently aware of their function. This unawareness works as an evidence filter and allows the acquisition procedure to tap the rich amount of subject inversion structures. Surface evidence for the directionality Subject-Predicate is due to the initial ignorance about the functional categories. If one considers the narrow frame of (22)a, shaped by adjacency plus theta assignment, the Subject-Predicate (68%) outnumbers the structure VSO (32%). This evidence is not overwhelming. Yet, the parameter is set almost immediately. Suppose, however, that not only the initial functional material Aux/Copula of (22)a is filtered out, but also the in between Aux/Copula of (22)b (Column C, now a+b combined). This yields the extended frame (22)a + (22)b, and the input evidence for the Subject-Predicate order raises to 73.5% versus 26.5% within the extended frame of relevance.

One might try to push the relative percentage over 80% by bringing in naming phrases (*dat is papa* 'that is daddy'), but the point should be clear. LF relevance frames, theta roles and predication, allow the learning procedure to filter certain evidence out of the input. Even modest assumptions show how input-control could in principle handle the Object-V and the Subject-Predicate directionality. The subject is now determined by lexical structure. The grammatical marking <+fin> has been left out by the theta/Case frame. This yields (23) as input evidence for argument order.

(23) Input evidence for argument order due to the theta/Case filter

Parameter	C ratio [+ evidence]	C ratio [- evidence]
Subject precedes predicate	73.5%	26.5%
Complement precedes predicate head	78%	22%

Let us now hypothesize that the acquisition procedure may set the argument order parameter to 100% if the decision is supported by some 70% of the distributional evidence. This type of parameter setting will often guarantee that somewhat messy input will nonetheless yield obligatory, exceptionalness rules.

This account of early language acquisition in terms of evidence filters proposed in Van Kampen (1997) and adopted here, is empirically sustainable. The two-word phase of early child language omits Auxes and Copulas and does so in spite of their high input frequency. Further, the two-word phase respects in general the Subject-Predicate and Object-Verb parameters. Finally, it may be observed that the scenario filters out the evidence for derived structure. It makes use of two factors, a factor of I-language (grammar) and a factor of E-language (input quantities). None of them represents a particular deep constraint. The 'filtering' I-factor is a blissful ignorance about the categories Aux, Copula and <+fin>. That suffices to separate the theta-assigning heads from their functional counterparts. The E-factor is the high percentage of Auxes in the input. These determine that most content verbs will appear in the predicate final position. As the amount of argument structures increases, it will strengthen Baker's positionally uniform theta assignment (UTAH) as a device to keep hold of the lexicon. This device will turn the VO structures from a minority pattern into an anomalous pattern. A situation that sets the stage for the introduction of the V-second rule, as we will describe in the next section. According to the present perspective, first language acquisition works only for those that do not have a full competence. Further, certain parameters can only be set if supported by broad quantitative proportions ('robust evidence') in the input i.e. E-language.[13] Neither a priori ignorance, that means lack of full competence, nor proportional weight in E-language look like an immediate favorite for generative grammarians. We add two remarks. Firstly, the input frequency of the functional items Aux/Copula dwarfs the frequency of any of the lexical content items, but their acquisition comes much later. Hence, the child does not parrot high frequency patterns that have some pragmatic benefit. Although controlled by input quantities, language acquisition by ignorance is a highly discretionary procedure. The success of language acquisition points towards contingent but felicitous learnability properties of I-language and E-language rather than to language specific a priori constraints of the neural system. Secondly, it is not claimed that predication is something that can be directly learned from input quantities. What is learned is a single form for a pre-existing naming phrase and a pre-existing characterizing phrase (cf. Van Kampen 2001). A non-linguistic pragmatic cognitive orientation is thereby grammatically schematized. The same sign, e.g. *bear*, *walk*, *on* ('the light/switch on'), can be used in both pragmatic modes, naming or characterizing. This is not an ambiguity between N and V. It is the initial absence of such distinctions. Lyons (1977, 1979) refers to this initial state of affairs as *proto-reference* (our 'naming function'), and *proto-predication* (our 'characterizing function').

[13] For learning from robust evidence, see also Valian (1990), Lightfoot (1991, 1999:156), Clark and Roberts (1993:330). Crucial reference to the category Aux for setting the directionality parameters is also present in Gibson and Wexler (1994:421) and in Fodor's reaction on the Gibson/Wexler paradox (1998:25,tab.1, 2001).

3.2 Apparent V-second

The preceding section showed how the input by the mother, characterized in (12), may reappear as a reasonable instruction about underlying argument order. The theta/Case frame from Lebeaux (1988) as interpreted in Van Kampen (1997), is to be applied as an evidence filter, see the table in (23). It has not yet been considered why that argument-verb filter does not get clogged by the overwhelming presence of finite verbs in the second position. Dutch as a V-second language invariably marks its root sentences by means of a full-fledged C-projection. Verbal as well as non-verbal predicates display this pattern, see the table in (24).[14] Table (24) is a further subdivision of (12)a in the preceding section. The percentage added to each type indicates its proportion in the input by the mother.

(24) Input percentages of <+fin> verbs by type
(mother Sarah, Van Kampen corpus)

Total <+fin> sentences			95%
Root <+fin> sentences			→ 93%
a. (XP)	**Aux**	[(YP) V]$_{VP}$	
b. (XP)	**Cop**	[Z (YP)]$_{ZP}$	→ 66%
c. (XP)	$\mathbf{V_{fin}}$	[(YP) t_V]$_{VP}$	27%
Subordinate <+fin> sentences			
d. ---	C^o	[(YP) $\mathbf{V_{fin}}$]$_{VP}$	→ 2%

Since 93% of the root sentences have a verb in the second position, one expects an effect on early child language. The distribution between <+fin> and <-fin> verb forms does indeed appear and it appears as early as the two-word phase. There is a further phenomenon to be observed. In as far as these forms occur, they are in accordance with the rules for root structures in the target grammar. This represents at the same time a more general state of affairs. For some deep reason, the child that picks up a grammatical device will hardly ever apply it mistakenly. The child may underspecify her/his utterances with respect to the target grammar, but s/he does so in a way that betrays a kind of awareness about the adult target construction. Before the child begins to utter syntactic forms, there seems to be a decisive orientation due to the rules and principles of the full system. This is the point of full competence views on language acquisition. Poeppel and Wexler (1993) and Borer and Rohrbacher (1997), among others, have argued that full competence offers the right perspective on the acquisition of Dutch/German verb distribution.

The full competence thesis is meant for a consciousness that perceives immediately how a priori UG principles apply to the grammatical structures. This might be like a visually equipped consciousness that turns its retinal impressions into a three dimensional orientation and shows no discernible problems in doing so. Like visual orientation, grammatical orientation could be due to domain specific turnkey devices of the human brain.

The language acquisition by grammatical hierarchy thesis is meant for a different kind of consciousness, a learning consciousness that shows considerable hesitance and evasion before it hits on the right combinatory principles. More crucially, the language acquisition by

[14] The following abbreviations have been used:

(i) a. XP, YP: arguments
b. Aux: finite auxiliary/modal
c. Cop: copula verb
d. Vfin: finite lexical verb
e. t_V: empty place of a theta-assigning verb

grammatical hierarchy thesis predicts that various acquisition delays can be related to inherent properties of the grammatical rule system. The logic of each thesis (full competence versus grammatical hierarchy) leads to a different analysis of evidence in the primary data.

3.2.1 V-second acquisition and learning by full competence

The full competence thesis is supported by an observation that, for some deep reason, is likely to hold quite generally. If the child acquires a grammatical device, s/he will hardly apply it mistakenly. For example, the use of <+fin> and <–fin> verbs is present as early as the two-word phase, and in as far as these forms appear it is in accordance with the rules of the target grammar. The correct use of the V-second pattern from the beginning on, is not immediately obvious, since children tend to leave out the finite forms {Aux<+fin>, Cop<+fin>, V<+fin>}. However, in principle competent use of the V-movement rule seems manifest, if one analyzes the primary data in a certain way.

The tabel in (25) below presents nine grammatical types in Dutch child language, {a-to-i} that have potential relevance for the acquisition of the V-second phenomenon. The relevance of these types is evaluated differently in three studies that are considered here. Two of them, Poeppel and Wexler (P.&W. 1993) and Rohrbacher and Vainikka (R.&V. 1995) are oriented by a full competence hypothesis, and one, Van Kampen (v.K. 1997) by the hypothesis that language acquisition proceeds in steps in accordance with grammatical hierarchy. The +, +* and – signs indicate whether the respective studies consider the type as, respectively, relevant evidence (+), counter-evidence (+*) or non-evidence (–) for the thesis that is under scrutiny.[15]

[15] Not all possible variants with topic-drop in <+fin> utterances are given in the table

(25) Relevant types for the acquisition of the V-second in
Rohrbacher & Vainikka (1995), Poeppel & Wexler (1993) and Van Kampen (1997)

			full competence		grammat. hierarchy
	Types	Examples	R.&V. (1995)	P.&W. (1993)	v.K. (1997)
a	**without Verb, no <+fin>** ('non-verbal predicates')	*papa daar/life/dokter/weg/op* (daddy there/nice/doctor/away/on) *papa boek* (daddy nice) *papa mee/aan* (daddy (go/come) with/on *daar papa* (there daddy)	–	–	+*
	with Verb, no <+fin>				
	two-word				
b	subj + V<–fin>	*papa lezen* (daddy read)	–	–	+*
c	non-subj + V<–fin>	*boek/ook/nu lezen* (book/also/now read)	–	+	+*
	>two-word				
d	subj + non-subj + V<–fin>	*papa boek/ook/nu lezen* (daddy book/also/now read)	+	+	+*
	with Verb, <+fin>				
	two-word				
e	subj + V<+fin>/Aux	*papa leest; ik kan* (daddy reads; I can)	–	–	+
	>two-word				
f	**type f = <+fin> opposite of type d**				
	subj + V<+fin> + non-subj	*papa leest boek/ook/nu* (daddy reads book/also/now)	+	+	+
	(non-subj +) V<+fin> + subj	*(boek/ook/nu/wat) leest papa* ((book/also/now/what) reads dad.)			
	subj + Aux (+ non-subj) + V<–fin>	*papa moet (boek/ook/nu) lezen* (dad must (book/also/now) read)			
g	**type g = <+fin> opposite of type a**				
	subj + Cop/Aux + non-subj	*papa is hier/lief/dokter/weg; p. heeft boek/moet mee/komt aan* (d. is here/nice/sweet/doctor/away; d. has book/must with/comes on)	+	+	+
	non-subj + Cop/Aux + subj	*daar is papa* (there is daddy)			
	ungrammatical order with Verb, no <+fin>				
h	V<–fin> non-final	*(papa) lezen boek/ook/nu*; *(boek) lezen papa* ((daddy) read book/also/now; (book) read daddy)	+*	+*	– (rare)
	with Verb, <+fin>				
i	(subj) + non-subj + V<+fin>	*(papa) boek leest* ((daddy book read)	+*	+*	– (rare)

The two studies oriented by the full competence thesis leave out type {a} in (25), *papa daar/life/dokter/weg/boek* ('daddy nice/doctor/away/book'). This is a considerable set in early child language, but type {a} contains no verb and hence cannot confirm or disconfirm a thesis about verb distribution.[16] Another irrelevance holds for the two-word utterances types {b/c/e} in (25). These cannot indicate a difference between the predicate final and the second position.

[16] Initially, the proportion of utterances with Verb is quite low, less than 10%. This is mainly due to the use of non-verbal predicates, type {a} in (25). See Behrens (1993:67,113).

Hence, they get a 'minus' sign. Poeppel and Wexler admit type {c}, nevertheless. They argue that this is a kind of three-word predication that opens with a dropped subject. The types {d/f/g/h/i} in (25) finally, allow a clear distinction between a verb in predicate final and a verb in second position. So all these types are relevant as (+). They support the thesis that child language from the beginning on brings the <+fin> verb in second position, as defined in the target language, whereas it keeps the <–fin> verb out of that position. If the types {h/i} had not appeared as marginal, they would have falsified the thesis of early full competence in V-distribution rules, and that defines their relevance as (+*).

Poeppel and Wexler (1993) considered the output of one day in the life of German Andreas (aged 2.1) in order to demonstrate how strongly he was oriented by full competence. They selected the >two-word types {c/d/f/g/h/i} for relevant evidence. These are the structures that contain a clear example of either a <+fin> verb or a <–fin> verb, or both. The selection was further determined by the question whether the structure showed indisputable evidence of a verb in the predicate final position or of a verb clearly not in the predicate final position. The resulting selection admits of the four possibilities in (26).

(26) Distribution of <+fin> and <–fin> verbs (Poeppel & Wexler 1993)

	<+ finite>	<– finite>
V-second	standard	rare
Verb final	rare	standard

The same method was applied by Rohrbacher and Vainikka (1995). The results of that study were used by Borer and Rohrbacher (1997) as further evidence for the full competence hypothesis. Rohrbacher and Vainikka looked at even younger children to find evidence for early V-second. They considered the data production of one recording session of Nicole (aged 1.8) and one of Katrin (aged 1.5). Again the thesis could be maintained. Child language respects rules for the distribution of <+fin> and <–fin> verbs from the very beginning. Decisive evidence for <+fin> elements in non-final position can only be derived from three-word utterances with a subject, a verb, and a non-subject (object/negative element etc…), but such examples are rare in early child language. In as far as they were present, they were seen by Rohrbacher and Vainikka as confirming the full competence hypothesis.

One may observe that this argument for full competence works as a verb-magnifier, since it leaves out the verb-less predicates of type {a} in (25). Moreover, in the language of the very young child, the two-word types with a verb are more numerous in the <–fin> variant, types {b/c}in (25), than they are in the <+fin> variant, type {e}in (25). Hence, the selection method works in addition as a <+fin>-magnifier. For example, Nicole (aged 1.8) and Katrin (aged 1.5) were for the most part in the two-word stage and Rohrbacher and Vainikka had to leave out 92% of Katrin's sentences with a <–fin> verb, i.e. types {b/c} in (25), but only 23% of the sentences with a <+fin> verb, i.e. type {e} in (25). For Nicole these percentages are 70% and 25% respectively. The highly focused selection should not diminish the awareness of the marginality that <+fin> verbs at this stage still have. This is a warning added to the percentages, not a criticism of the method. One may even argue that the correct <+fin> distribution is even more striking now that it holds for a quantitatively marginal set of predications. Neither an explanation by habit formation, nor an explanation as frozen formula is easily compatible with marginality.

3.2.2 Apparent V-second and learning by grammatical hierarchy

Language acquisition by grammatical hierarchy implies that the child selects the input data in a different way. The initial grammar of the young speaker forces him to leave out the functional categories <+aux> and <+copula>. This affects the intake of the predicate constructions in (24)a/b, i.e. 66% of the input predications, and predicts the appearance of the types {a/b/c/d} in (25), see the examples repeated in (27).

(27)	a. papa daar/lief/dokter/weg/op	(daddy there/nice/doctor/away/on)
	papa boek	(daddy nice)
	papa mee/aan	(daddy (go/come) with/on
	daar papa	(there daddy)
	b. papa lezen	(daddy read)
	c. boek/ook/nu lezen	(book/also/now read)
	d. papa boek/ook/nu lezen	(daddy book/also/now read)

These types appear in the input by mother by exception only. Yet, they are the norm, rather than the exception, in early child language. The systematic ignoring of <+aux> material does not suffice to explain why the initial output by the child also ignores lexical finite verbs in the second position (type (24)c). The input predications by the mother offer such types for more than a quarter of the total set of predicates, cf. (24)c and (25)e/f. see the examples in (28).

(28)	a. papa leest boek/ook/nu	(daddy reads book/also/now)
	b. papa leest	(daddy reads)
	c. boek/ook/nu/wat leest papa	(book/also/now/what reads daddy)

These predications are theta assignors, and hence they are expected to be relevant intake. By the logic of the present analysis, the acquisition procedure should process type (28)a as an alternative theta frame and not as a movement structure. Its appearance in the initial production of the child is marginal to non-existent. At first less than 1% of the predications, and later on some 10%, have a lexical V<+fin>, as we will show in the next section. The full competence thesis was grateful for each example of a finite verb in the child's output. The present thesis is different. If the adult input of V<+fin>-Object structures appears to the child as an alternative theta frame, which covers 22% of the input (cf. (21)a), one expects the same percentage in the output of the child. This expectation is simply not met and it requires an explanation. It cannot be claimed that the child evades grammatical markings in general and that not only auxiliaries and modals, but also lexical finite verbs fall under that restriction. By the logic of the present hypothesis there is no <+fin> as a grammatical marker yet. There is only a theta-assignor. Suppose the following subplot. Quantitative evidence leads the acquisition procedure almost instantaneously to the OV parameter, cf. 78% over 22 % in table (21). It might be that the acquisition procedure starts a subroutine and after some delay introduces an alternative frame V<+fin>-Object. This would cover 22% of the input as offered by the mother. If so, the initial attention of the child to the morphological opposition <+fin>/<-fin> should relate to a switch in the direction of the verb complement theta frame rather than to the (later) movement rule.[17]

Even if the acquisition procedure would permit such alternative frame V<+fin>-Object, there is an almost exclusive preference by the child for the major pattern OV. The <+fin> forms are highly marginal before the appearance of the periphrastic forms. The low percentage

[17] See Van Kampen (1997: chap.2; 2001) for a comparison between the present approach and Braine's pivot grammar.

is especially telling, since there is very little overlap such that predicate heads <+fin> also appear with <-fin> variants in the predicate final position. Such overlap does not appear until the rise of periphrastic forms with a <+aux/+fin>….V<-fin> form is well on its way (Wijnen 1997, Van Kampen 2001). See Blom (2002) for a careful quantification of early overlap.[18] Personally, we would prefer to push aside all early appearances of lexical V<+fin>. As initially less than 1% of the predications, and less than 10% later on, they seem too marginal to pay attention to. The subroutine is offered for those that would like to account for early lexical V<+fin> anyway.

3.3 Longitudinal graphs

Quantitative data from the longitudinal development support the view that child language reflects a succession of grammars that access a grammatical hierarchy. A series of spontaneous conversations between the young speaker and the adult will contain a number of predicative constructions. The percentage of <+fin> marked predications can be determined for each conversation. The construction of a graph proceeds by setting out the percentages on a time scale. The 100% line in (29) represents the full set of predicational structures for each measuring point. It considers all multi-word predicates in the files of that period and divides them in a part characterized by a <+fin> element in first or second position, i.e. types {e/f/g} in (25), and a part not characterized that way, i.e. types {a/b/c/d}. In this way, the graphs in (29) depict the acquisition of the V-second structures by Laura and Sarah (CHILDES, Van Kampen corpus).[19] The measuring points of (29) begin after week 100, after the rise of Sarah's and Laura's two-word predications. Measurements before that point get highly dependent on arbitrary decisions about one-word utterances as predications. For example, is a predication like 'bear!' ('that is my bear' or 'I want my bear') a predicate that lacks a <+fin> marking? (cf. Van Kampen 2001). Yet, the one-word utterances are an overwhelming characteristic of that period.

[18] Blom (2002) begins with Wijnen's (1998) assumption that the early <+fin> patterns are basically a semantic class <+stative>. She ends with the position that they are merely syntactic patterns imposed by frequency as argued in Van Kampen (1997).

[19] The graphs in (29) are based on the following number of relevant sentences (types {a-to-g} in (25)), disregarding for a moment their asymmetric distribution in time.

	total number	V <–fin> types b/c/d	V<+fin> types e/f/g	'non-verbal predicates' type a
Laura	3.537	651	2.532	354
Sarah	2.789	282	2.181	326

(29) The acquisition of V-second

Comparable graphs for the acquisition of V-second have been constructed earlier by Ruhland, Wijnen and Van Geert (1995) and Wijnen (1997) for Peter (aged 1;9.6-2;4.12) and Matthijs (aged 1;11.10-2;11.19). These studies paid much attention to the question whether the measurements are sufficiently fitting a logarithmic curve.[20] Our concern is the more linguistic question whether this type of measurement can be used to discover the robustness and the ordering of the various learning steps. The straight line in (29) between the last time that the

[20] See also Ruhland (1998:chapt.5.2), who concluded that a logarithmic S-curve yields a closer approximation of the measurements than a three part linear graph does. More complex functions could in turn outshine the logarithmic curves. This is interesting only in so far as more mathematical sophistication in the graph leads to actual conclusions about the acquisition of grammar. For our present concerns (order and speed of parameter setting), three-part linear graphs are sufficiently revealing.

curve crosses the 10% level and the last time it crosses the 90% level, can be used as a delimitation of the learning step and a measurement of its speed or robustness. Quantitative data from the longitudinal development support this view. The crossing points themselves at the 10% and the 90% line can be indicated as the eureka point and the acquisition point, respectively. If the eureka point and the acquisition point are the same moment in time, one may say that the acquisition is instantaneous. If the eureka point of one grammatical change follows the acquisition point of another grammatical change, one may say that the learning steps are ordered. Thereafter all these types of facts about the order and speed of acquisition steps are listed as problems that require an explanation.

The acquisition point in (29) is constituted by the measurement where the <+fin> predications cross the 85% line. At this point the child is within 10% of the adult norm (i.e. 90% of the >90% V-second in adult Dutch). Sarah and Laura reach the acquisition point at different moments, but the steepness of the acquisition lines is rather uniform. Sarah's percentages of predications that is marked by V-second is rising for some 20 weeks from below 10% to more than 85%. Laura's case fits this pattern if one disregards the horizontal line half way, where her development stagnates for some 30 weeks at 55%.[21] If one could abstract from the stagnation period, Laura would also fit a normal 20 weeks picture (4.5 + 15 weeks).

The more steep the acquisition line is, the more speedy the acquisition step and the more robust the grammatical rule that is being acquired. The development between eureka point and acquisition point must be a matter of getting the rule speedy/automatic, such that it can be applied in running conversation. One might label the time span between seeing the rule (eureka point) and getting it automatic (acquisition point) as 'performance delay'. The period before the eureka point may be labeled as 'perception delay'. It is a point of potential interest that children might have considerable differences in their perception delay, whereas the performance delay is rather uniform.

The rising percentage of V-second structures in (29) is partly based on finite Auxes and Copulas, patterns (24)a and (24)b, and partly on the finite form of lexical content verbs, pattern (24)c. The latter type is somewhat slower to take up its 30% share in the <+fin> marking of the input. This is demonstrated in the columns of (30).

(30) <+fin> Aux/Vlexical columns Laura and Sarah

<+fin>		S1	S2	S3	M	L1	L2a	L2b	L3	M
Aux(-V)	number	86 n	140 n	297 n	731 n	72 n	155 n	156 n	242 n	690 n
	%	92%	80%	70%	72%	94%	87%	88%	73%	69%
Lexical V	number	7 n	36 n	127 n	283 n	5 n	18 n	22 n	88 n	303 n
	%	8%	20%	30%	28%	6 %	13%	12%	27%	31%

The columns S1, S2, S3 present Sarah's <+fin> marking by means of functional verbs and lexical content verbs in the beginning, the middle and the end of the 20 weeks period of the V-second curve.[22] The columns L1, L2a, L2b, L4 represent Laura's acquisition of <+fin> marking. The columns L2a and L2b reflect Laura's V-second system at the beginning and the end of the 30 weeks stagnation period. The graph in (29) and the columns in (30) for Laura

[21] Laura's delay in developing the V-second structures is probably due to repeated ear infections. Unfortunately, no notes have been made of her sicknesses.

[22] The following files were selected. For Sarah: files 9,10: 86n-7n=92%-8%; 11-13: 140n-36n=80%-20%; 17-19: 297-127=70%-30%. For Laura: files 11-13: 72n-5n=94%-6%; files 17-19: 155n-18n =87%-13%; files 22-24: 156n-22n =88%-12%; files 31-33: 242n-88n = 73%-27%. Imperatives and subordinates were left out. Of course, the percentages are based on a lower number of utterances for the first files than the percentages for later files.

show that during this period there was no rise in <+fin> structures, nor in the number of lexical V. A trend-setting column M (mother), set on some 1000 consecutive V-second utterances, has been added for each child. Column M then represents the (24)a/(24)c relation of the adult input. If more material had been available, separate curves for functional and lexical <+fin> marking might have been constructed. The acquisition of the lexical <+fin> elements is somewhat different and obviously more difficult. The files show that constructions with a <+fin> marking are present from the beginning, but partly as fixed expressions only, and marginal anyway, <10% as compared to the 90% <–fin> and <fin>-less constructions. When the amount of <+fin> predications has finally begun its irreversible rise, the >90% line is reached within 15 to 20 weeks. The columns in (30) show the delayed rise of the minor pattern (24)c. The rule 'move $V_{<+fin>}$' realizes at first only a quarter of its later share in predicate marking and remains quantitatively less prominent until the acquisition point has been reached. Nevertheless, movement pattern (24)c is eventually acquired.

It is only after the acquisition point of the V-second rule that the first subordinate structures appear. They remain initially rare, some 1-2% of all predications, as in the adult input at that moment. There are 20 examples in the files of Laura between week 153-198 and 36 in the files of Sarah between week 135-183. The finite verb, that used to appear in C^o, turns up now in the predicate final position. This is a dramatic distributional difference with the root clause, as pointed out in section 2. If one considers the first two-word utterances as a syntactic starting point, it took the language acquisition procedure some 20 weeks to reach the eureka point (19 weeks for Laura, 21 weeks for Sarah). Thereafter it took another 20 weeks to reach the acquisition point. Since nearly all input predications are characterized by the V-second rule, the evidence for the rule must have been massive. For a period of 300 days and a few thousand examples of input V-second structures a day, more than half a million examples were needed to adapt the child's system and to reach acquisition point. The distribution of verbs in the subordinate is rather different from the distribution in the root clause. Nevertheless, the acquisition of the subordinate pattern is instantaneous. All examples of subordinates appeared just after the V-second acquisition point, and none of them failed to have the correct verbal distribution. This difference in acquisition achievements needs an explanation.

In sum, four major longitudinal facts have been singled out.

(31) a. The acquisition procedure initially disregards the <+fin> marking for almost half a year (20 weeks), although <+fin> is present in almost any predication of the input (95%).

b. A grammar with <+fin> marking eventually blocks an original and more simple grammar without <+fin> within, again, 20 weeks.

c. The acquisition of finite lexical content verbs, pattern (24)c, trails the acquisition of finite Auxes, patterns (24)a/b, for some reason.

d. The acquisition of the verbal placement in subordinate structures is instantaneous, although it is a rare and confusingly different pattern of distribution.

These four facts fit the incremental acquisition that follows from the grammatical hierarchy hypothesis as will be argued below.

3.4 Reconstruction of the verbal chain

The acquisition procedure manifests an initial disregard of <+fin> marking, cf. (31)a. The input marks 95% of its predicates by <+fin>. Early child language typically lacks these characteristics. This is easily explained by the grammatical hierarchy hypothesis, cf. Van Kampen (1997:chapt.2). Any <+fin> marking applies to a lexical structure. The <+fin>

marking cannot be perceived or learned before the lexical structure it applies to is readily available. The acquisition procedure by grammatical hierarchy is forced to first scan the input for lexical frames, as assumed above section 3.1, see (32). Such an approach is bound to leave out the <+fin/+aux> markings and, as a matter of fact, they are left out.

(32) Lexical structure $<_{\text{precedes}}$ Functional structure

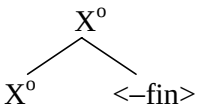
For the primacy of lexical structures see Lebeaux (1988) and Radford (1990), and many others. Acquisition by grammatical hierarchy leads to the correct prediction that the patterns (24)a and (24)b will appear as types {a/b/c/d} in (25), forms that are nowhere in the input. The highly frequent functional categories Aux and Copula are left out. A second prediction is made. The finite lexical verb, pattern (24)c, cannot be perceived as the result of a V-second movement. The V-second rule in the target grammar moves the lexical predicate head from its base position into the functional position C^o. As long as no functional position has been constructed, no such movement can be identified and learned. Acquisition by grammatical hierarchy forces the conclusion that pattern (24)c is at first perceived and learned as an alternative theta frame, as in (33).

(33) a. YP + X<–fin>
b. X<+fin> + YP

As one can see in (21)a, pattern (33)a is supported by 15% of the input predications and pattern (33)b by 5%. Hence, there must be a strong preference for (33)a (relative 75%-25%). Two different theta-assignors, X^o<–fin> and X^o<+fin>, are introduced. There is as yet no paradigm {X<–fin>, X<+fin>}. These two forms are merely lexical look-alikes, each base-generated with its own theta frame. The initial interpretation of early <+fin> as no more than a lexical feature is also present in Wijnen (2000:168). One might compare this with a proposal for the adult language where all or certain passives are base-generated from lexical variants of the verbal head. (Levin and Rappaport 1986). As far as the <+fin> forms in child language are concerned, Sarah and Laura used pattern (33)b at the eureka point for no more than 1% of all predications (see the columns in (30)). A careful count by Blom (2002) revealed that 'overlap' is initially rare. There are hardly verbs that appear in both variants <+fin> in non-final and <-fin> in final position.

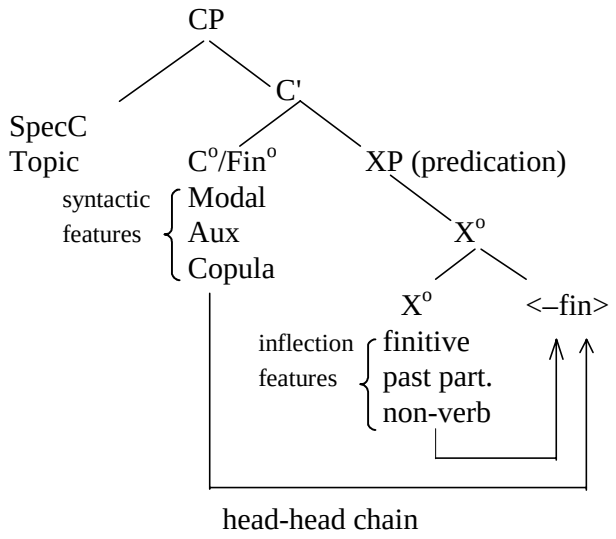
The acquisition of the UTAH theta-frames sets the scene for the second step, see (31)b. In the adult input, the predications will not appear without a <+fin> marking. The <+fin> marking has two functions. Semantically it expresses the illocutionary value of the predication and syntactically it licenses the form of the predicate head. Certain predicate heads require certain <+fin> elements as their licensers. There are language specific head-head chains to be perceived and learned. The X^o in the predicate final position can be an infinitive, *te*+infinitive, past participle, or non-verb. Each of these elements requires a specific <+fin> element in C^o. There are a lot of details here, but it is for sure that the insertion of <+fin> items cannot take place unless these various co-occurrence relations have been perceived and acquired. The co-occurrence relations between the <+fin> element in C^o and the X^o head in predicate final position can be modeled grammatically. The two heads have to be co-indexed. Assume that the lexical content element X^o in the predicate final position is indexed by the feature <–fin>, and also that <–fin> is being adjoined to the X^o as in (34).

(34)



The <+fin> stands for a set of syntactic features and the minus sign stands for 'needs licensing by' that set. The chain index <–fin> is a clear interface feature. It has to answer syntactic conditions within the extended XP as well as morphological conditions within the head X°. The <–fin> index can be removed from the PF representations if it matches with the syntactic <+fin> in C°, and with the morpho-syntactic choices made by the X° item. In diagram:

(35)



A crucial point has to be made here. The introduction of the <-fin> index implies that the lexical heads X° in the predicate final position get a new categorial status. They change from neutral self-sufficient X° theta-assignors into a few different categories {*te*+infinitive/bare infinitive/past participle/rest-category}. These new categories are all structure dependent in a head-head chain. As long as the lexicon contained the neutral X° heads, the <+fin> marking of predicates was not called for. As soon as the lexical head X° tends to be interpreted as [X° + <–fin>], it requires syntactic licensing and the <+fin> element must be present in syntax.[23]

It is not assumed that the functional categories summarized here as Fin° represent a tense or an agreement marker. Agreement will remain deficient for some time, as in all non pro-drop languages. Correct and systematic agreement is mastered long after the acquisition of the verbal positions, see Schlichting (1996), Van Kampen and Wijnen (2000). Nor does Fin represent a present tense. All predications can be interpreted as present tense by an adult grammar, but an adult grammar is bound to over-interpret. Early child grammar offers no tense oppositions. For that reason, it offers, grammatically speaking, no tense. The oppositions present within the Fin of child language are best characterized as simple, inflection-less, modal and aspectual operators like {assertion/command/request; state/event}. This point is due to De Haan (1987).

The <–fin> index of the X° in the verbal chain is assumed to attract all grammatical information from the Fin° head of the chain. This proposal is confirmed by a strange fact. The acquisition of the subordinates structures is instantaneous, see (31)d. The subordinate structure

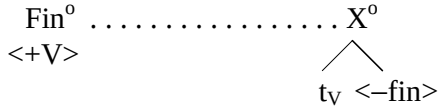
[23] The category <+V> make no sense in the grammar of Dutch before the <+fin>/<–fin> mechanism is in place, see Van Kampen (1997:chapt.2). It may be maintained that the feature <+V> merely restates that an X° has access to a <+fin>/<–fin> paradigm.

has its C^o position marked by the constant *dat* ('that'). This might be interpreted as signal for an <–illocution> sentence. The <–fin> X^o elements in predicate final position (infinitive, past participles and rest-category) appear but according to their new categorial status they will continue to require licensing by a head Fin^o. The head Fin^o now appears in the position of the <-fin> index. By the analysis above this position did already contain all syntactic information of the Fin^o. Only the illocutionary effects are now blocked by *dat* 'that (<-illocution>). An interpretation of the subordinate as a simplified and de-illocutionized variant of the root structure has been argued for by Den Besten (1977/1983). Den Besten's (1977/1983) leading idea may suggest that the subordinate structures in most languages will be acquired instantaneously. It might demonstrate by a quite general argument that early language acquisition must be a matter of rules and principles, rather than an assimilation to frequently heard patterns, contradicting Tomasello (1999). The present analysis constructs a head chain and predicts that the subordinates will not appear until the V-second rule and its verbal chain have been fully apprehended. Before the full apprehension of the verbal chain, the subordinates must impress the young language learner as a temporary gush of distributional garbage. After the construction and full acquisition of the verbal chain, the subordinate appears as a simplified variant of the root structure. Its syntactic form has already been acquired as a subpart of the root structure. It is the XP in root structure (35) above. Its recognition as a subpart makes that the eureka moment of the subordinate coincides with its acquisition moment, as stated in (31)d.

The proposal that subordinates are acquired after and due to a previous acquisition of the root structure contradicts persistent suggestions in the literature that the opposite must be true, at least for the case of Dutch and German. According to these suggestions, awareness of the subordinate structure would help the analysis of the root sentence (Roeper and Weissenborn 1990, Penner 1993) Quantitatively, this conjecture is not convincing as 2% of the input is subordinate (see (24)d). More important is that the suggestion suffers from an inconsistent acquisition logic, as already pointed out in Van Kampen (1997:41). The criticism runs this way. Suppose, one would argue that quantitative input properties are not that important. There is input enough. Even 2% is a lot of structures. Moreover, as an a priori minded colleague observed, what to think of those toddlers? Are they counting all the time and do they keep an eye on percentages and proportions like a bunch of players on the stock market? Suppose we accept this criticism and remove the 70% input requirement on distributional intake filters. Quantities do no longer matter that much. Now there arises conflicting evidence all over the place. There are infinitival <-fin> and finite <+fin> predications. The latter group <+fin> contains predications with a predicate final <+fin> next to predications with a predicate initial <+fin>. One might now propose boldly that the child is a full competence learner. S/he knows about Ross' Penthouse Principle (the downstairs, i.e. the subordinate, structures show the base conveniences only). Hence, s/he is smart enough to concentrate on the subordinate structures, somewhat more rare though they may be. Let us grant that as well. It won't help. The Gibson/Wexler paradox comes now in with full force. The young learner cannot find out which structures are subordinate, unless s/he is attentive to subordinate marking as such, verbal distribution for example. Yet, these very distributional patterns are the ones the child does not know. They are language specific and by assumption had to be learned. The Gibson/Wexler paradox has incisive force and is relevant to all present day acquisition literature. Although it has been proposed that the underlying verbal order in Dutch/German is attainable for an acquisition procedure attentive to subordinate construction (Roeper and Weissenborn 1990, Penner 1993, Powers and Lebeaux 1998) or attentive to negation (Lebeaux 1988) or attentive to particle distribution (Lightfoot 1991) or attentive to stress patterns in the subordinate (Nespor et. al 1996), none of these works deals with the question how such hints can be part of the intake.

This leaves us with the statement in (31)c. Why is the shorter and apparently more simple pattern (24)c that contains no auxiliary or copula acquired later than the patterns (24)a and (24)b. Clearly, the verbal chain of real PF heads in Aux and Copula constructions, patterns (24)a and (24)b, does not immediately stand out in structures like (24)c. There is in (24)c a finite lexical head in C^o and no head in the predicate final position. It stands to reason that the 27% predication structures (24)c in the input call some attention due to their slightly anomalous form. Assume that the preliminary grammar tends to expect that these predications should have a theta-assignor in the predicate final position as well, since that would confirm the more frequent and the more dominant UTAH pattern. That expectation is not met by the input structure of pattern (24)c. Suppose we model this unfulfilled expectation grammatically by an empty head X^o, see (34) This empty head is grammatically chain-related to the finite (lexical) head in the second position by the same device already in use for (24)a and (24)b.

(36)



The delay of V-second for lexical verbs, pattern (24)c, seems more reasonable now. The pattern (24)c does not fit the OV UTAH pattern. It does not fit the V-second rule and its previous construction of a verbal chain. Two learning steps must be taken to reach that point. Firstly, the missing predicate head must be modeled as t_V. Thereafter, the <+fin/–fin> head chain must be perceived and acquired for pattern (24)c.

The conjecture that pattern (24)c is acquired after and due to the acquisition of pattern (24)a was made by De Haan (1987). It has been disputed by Poeppel and Wexler (1993). The latter objected that German Andreas (aged 2.1) acquired a regular V-second grammar quite early. Our evidence indicates that this as such is no argument yet. One needs longitudinal arguments. The V-second acquisition graph is fairly steep. Recordings of no more than ten weeks earlier might have shown that pattern (24)c was still anomalous or marginal.

It has been argued in this section that language acquisition by grammatical hierarchy can plausibly reinterpret early V-second facts brought forward by the full competence hypothesis. In addition four longitudinal phenomena in (31) have been brought in perspective. The acquisition of V-second and the elaboration of the <+fin>/<–fin> paradigm in the lexicon has been related to the construction of a verbal chain. The grammatical features that characterize dependency in the chain are language specific. Favorite sons of UG and a priori full competence like <+tense> and the pronominal distinctions of <+agr> play no role at all. They are filled in much later. A full competence hypothesis may be in a more difficult position to explain the longitudinal facts in (31). The better the full competence hypothesis explains (31)d, the instantaneous acquisition of the subordinates, the more it is in trouble with an account for (31)a and (31)c. Why is there according to a full competence hypothesis an initial delay in the spell-out of predicate licensing? Why do the categories of Tense and Number play no part if that licensing finally comes in and why does the delay in predicate licensing extend to type (24)c, point (31)c? Finally, how does the full competence hypothesis explain that the root infinitives get blocked eventually, cf. (31)b?

One might construe the disappearance of optional root infinitives in two different ways. The full competence approach perceives predicate VP structures and these are always licensed. The spell-out of the licensing is optional in the beginning, see the scheme in (37).

(37)

	Gi:	optional spell-out of licensing
	Gi+1:	obligatory spell-out of licensing
	Gi $\longrightarrow$	Gi+1 the acquisition moves from a superset (Gi) to a subset (Gi+1)

This superset situation might hold for the acquisition of any adult licensing principle in a full competence approach. The approach by grammatical hierarchy perceives predicates, VP structures among them, and these are initially never marked by licensing material. The appearance of licensing material involves as well a reinterpretation of the previously unlicensed predicates as licensees, see the scheme in (38).

(38) Gi: predicates unlicensed (no finite verbs, no copula)
Gi+1: predicates licensed
Gi ⟶ Gi+1 moves from one kind of grammar (no licensing) to another kind of grammar (licensors and licensees)

There is no superset/ subset relation for the acquisition of adult verbal licensing in the present analysis.

3.5 V-second $<_{precedes}$ wh-movement

Content questions appear quite early in Dutch child language, but their marking by means of an initial wh-constituent does not take place before the V-second property is present, consider the patterns in (39). The form (39)a, equivalent of the question *wat doe je?* ('what are you doing?'), appears marginally and as fixed formula during the period of 'optional infinitives'. Curiously, type (39)b does not appear. Type (39)d is the adult target form, with the full fledged C-projection. One might have expected that the (39)d types would appear simultaneous with or immediately after the rise of V-second construction, i.e. after the disappearance of the 'optional infinitive', but this is not the case. Type (39)d is preceded by type (39)c, a form which is not or hardly present in the input, but quite characteristic for a period in child language. .

(39)	a. ---	---	jij nou	---	doen ?	'optional infinitive'	(∅ you now do<–fin>?)
	b. wat	---	jij (nou)	---	doen ?	'not attested'	(what you do<–fin>?)
	c. ---	doe	jij nou	---	?	'precursor'	(∅ do<+fin> you now ?)
	d. wat	doe	jij (nou)	---	?	'adult target'	(what do<+fin> you?)

Leaving the Spec,C unfilled in adult Dutch marks the sentence as a yes/no-question. If simultaneously one of the arguments is left out, there is a double gap construction (39)c that is used by the children as a content question. Type (39)c is typical in child Dutch and not justified by any direct parental input. The mother's input sentences in the files left out the wh-marking of content questions in less than 2% of the cases, whereas the children started with a period of >90% wh-drop.[24] The precursor form (39)c evades a functional category <+wh> and probably a case of explicit movement as well. The previously acquired finite verb in C^o serves as a reliable identification for the underspecified and phonologically empty Spec,C. Later we will introduce the notation <F?> for the initial emptiness of grammatically designated categories. The fact that there is hardly any support for (39)c in the language of the mother is

[24] The 2% adult wh-drop in Dutch is restricted to informal fixed phrases. An example of the mother of Sarah is given in (i).
(i) doe je nou, Saar? ((what) do you then, Saar? = what are you doing, Saar?)

The children apply wh-drop for all predicates, see some examples in (ii), taken from Van Kampen (1997).
(ii) heet zij nou? (calls she then? = (how) does she call/what is her name?)
is deze voor nou?(is this for then? = (whom) is this for?)
ga jij nou heen? (go you then to? = (where) are you going?)

quite remarkable. It has been argued that (39)c must be considered as an extended application of topic-drop in Spec,C by means of an underspecified A-bar pronoun (Van Kampen 1997:chapt.4).[25]

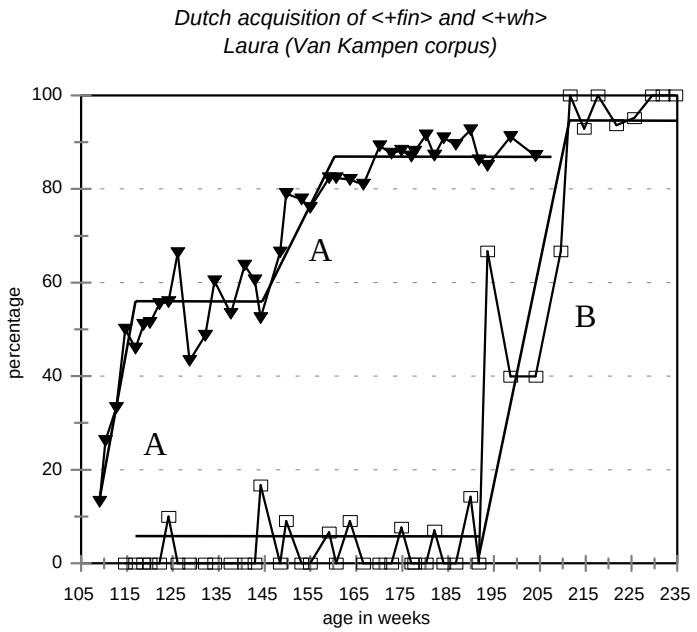
The correct use of wh-pronouns is not beyond the reach of the young speaker's capacities. At first, somewhat less than 10% of the content questions (with a <+fin> marking) are marked by the initial wh-pronoun, most of the time a fixed formula. Nevertheless, it takes some time before the percentage of wh-questions within the set of content questions starts to rise. Once the percentage crosses the 10% line the development seems irreversible. Within some 15-20 weeks the 90% line is reached. See for this development the graphs of Laura and Sarah in (40).[26]

[25] The missing wh-elements concern pronominal information of lexically selected elements. Therefore the argument gap can be identified by means of the theta-grid of the verb. See Van Kampen (1997) for a quantified analysis.

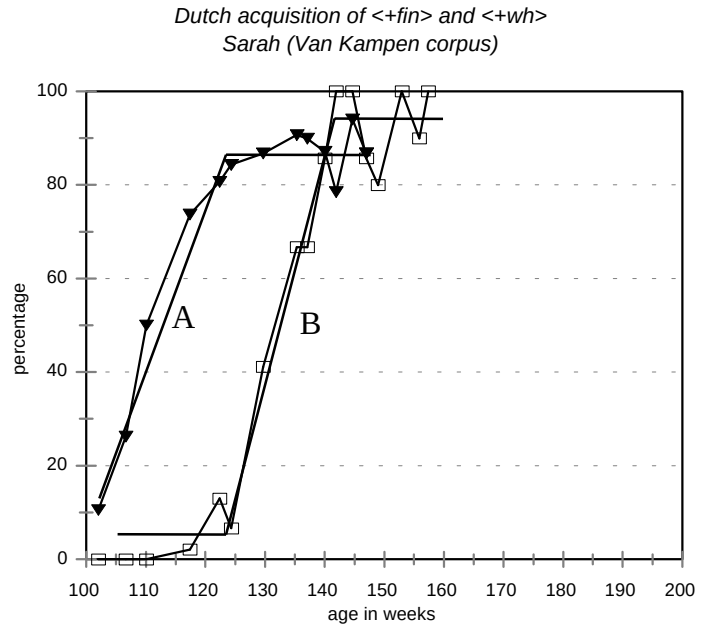
[26] The graphs in (40) are based on the following numbers, disregarding for a moment their asymmetric distribution in time.

	Total number of wh-questions	with double gap	with a wh-element
Laura	491	228	263
Sarah	497	127	370

(40) Dutch CP: <+fin> $<_{precedes}$ <+wh>



Graph A: Sentences that realize <+fin> in first or second position: **4.5+15 weeks=19.5 weeks**
Graph B: Content questions that realize <+wh>: **20 weeks**
Acquisition point A → Eureka point B: **33 weeks**



Graph A: Sentences that realize <+fin> in first or second position: **20 weeks**
Graph B: Content questions that realize <+wh>: **17.5 weeks**
Acquisition point A → Eureka point B: **3 weeks**

The acquisition of wh-marking in content questions is remarkably similar for both children. Laura and Sarah do not introduce the wh-pronouns as fixed elements before the V-second rule has been acquired. Both children have a period in which content questions are present while the wh-marking of them holds for a marginal subset only. Both children also show the irreversible 15-20 weeks growth of the wh-subset towards a 100% obligatory wh-marking.

There is a striking difference as well. Laura suffered during early childhood from heavy colds and repeated ear infections. The language files that were analyzed five years later showed how she trailed Sarah's grammatical development by more than a year. Nevertheless, once her switch to the wh-subset started, she made it within the same 15-20 weeks.

This suggests that the eureka point, the point of definite grammatical insight, is more a matter of chance. Its moment of appearance might be more or less fortuitous. The further development from eureka point could be a far more mechanical effect of habituation with a more predictable timetable. The new rule is not a part of the current grammar until the acquisition point has been reached. From that point on it is a part of the current grammar and possibly an essential support for reaching the next eureka point.

This might be a good point to return to the Gibson/Wexler paradox, referred to in section 1. The following thesis for the acquisition of the C-projection has been constructed. Firstly, the lexicon is set up as a uniform array of fixed argument frames, with the verb in predicate-final position (Baker's UTAH). The possibility of this learning step has been demonstrated by Brent (1994). Moreover, it fits and even explains Baker's UTAH for underlying structure. It sets the C-projections apart. The child ignores them (<+wh>) or accepts them only as marginal possibilities (<+fin>). Secondly, functional categories such as <+fin> and <+wh> become interpretable as soon as, but not before, there is an ordered argument structure of lexical content items that functional categories can apply to. This view on functional categories derives the order of learning steps from grammatical hierarchy and it solves the Gibson and Wexler (1994) paradox about the inaccessibility of underlying directionalities.

4 Order of parameter setting in English

4.1 The Dutch-English differences

As Dutch, English can express the illocutionary values of a predication by an X-bar envelope around the IP/VP, known as the C-projection. Both languages build up the C-projection by extracting material out of the IP/VP. The use of the C-projection in English is more restricted than it is in Dutch, see section 2. All root structures in Dutch are marked by the C-projection, whereas in English only root questions are. Moreover, English moves only <+fin,+aux> elements into the C^o position, whereas Dutch brings in <+fin,–aux> elements as well. See diagram (41).

(41)

Root clauses	Spec C	C^o <+fin>
Dutch	+/– wh	+/– aux
English	+ wh	+ aux

The acquisition of wh-questions in English differs greatly from the acquisition in Dutch. Dutch child language starts the marking of constituent questions by inserting the <+fin> verb in C^o and by leaving the Spec,C unfilled. The English acquisition order is reversed. The Spec C is acquired before the C^o is filled in. The acquisition periods are quite different as well.

(42)

Dutch:	move <+fin,+/–aux> to C^o *15-20 weeks*		$<_{precedes}$	move <+wh> *15-20 weeks*
English:	move <+wh> *instantaneous*		$<_{precedes}$	move <+fin,+aux> to C^o *a year (50 weeks)*

This will be demonstrated by longitudinal graphs based on the Brown corpus. Section 4.2 deals with the acquisition order between move <+wh> and move <+fin,+aux> in English child language. Section 4.3 considers a special acquisition difficulty of English. In order to spot the rule 'move <+fin,+aux> from I^o to C^o', the English child must first construe a chain C^o-I^o. Again, it will be argued that underlying structure (in this case the chain C^o-I^o) can be derived from quantitative proportions in the PF input.[27]

4.1.1 English wh-movement $<_{precedes}$ residual V-second

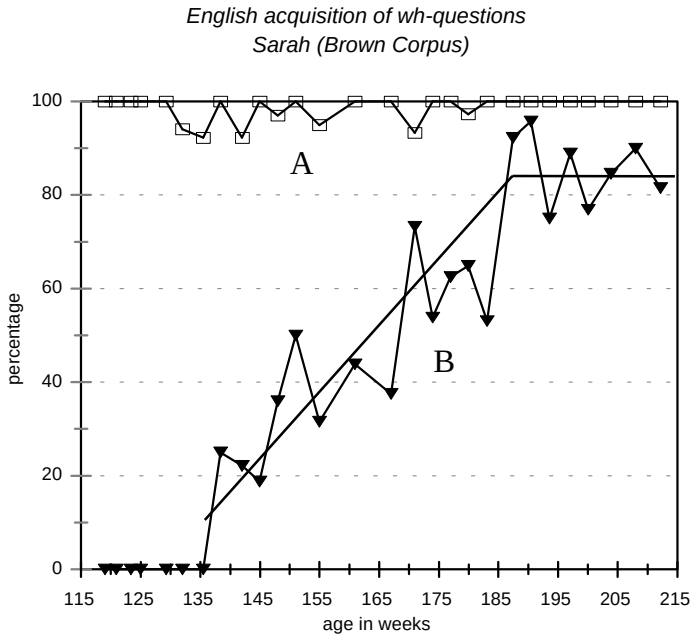
The wh-construction for root questions in English must have a wh-phrase in first position. If the wh-phrase is a non-subject, it is in Spec,C and it is invariably followed by a <+fin,+aux> element. The <+fin,+aux> element is most of the time cliticized on the wh-phrase and not particularly prominent. In the case of one child, the percentage of <+fin,+aux> cliticized on the wh-element in the speech of the English mother was as high as 77%.[28] Subject questions have a wh-phrase in the subject position (Spec,I) and do not need the use of a <+fin,+aux> head. Hence, the invariable property of the wh-root question is the wh-phrase and not the (cliticized) <+aux> element. It now appears that the presence of <+fin,+aux> in the C^o of English root questions is difficult to acquire. English Sarah (CHILDES, Brown corpus) took a full year, as is shown by the longitudinal graph B in (43).[29] The measuring points in graph B are based upon the obligatory presence of <+fin,+aux> in the C^o position of non-subject wh-questions. The (shrinking) percentage of finite auxiliaries lacking in the child's output shows to which extent the auxiliary system still functions as an optional marking. Graph B is somewhat unruly, probably due to the fact that the measuring points contain an insufficient amount of cases (between 12 and 35), and to the fact that presence or absence of a cliticized Aux is not always easy to establish by the transcriber.

[27] See for an input-driven account of the acquisition of subject-aux inversion in English also Weinberg (1990).

[28] We counted 493 wh-questions used by the mother of Sarah (CHILDES, Brown corpus) in the files 1-17, 20 subject and 473 non-subject wh-questions. Of these 493 wh-questions 380 (77%) had an Aux cliticized on the wh-element.

[29] The graphs for English Sarah are constructed from the Brown corpus in the CHILDES archive, from files 1-90, recording Sarah's first acquisition of English between 2;3.5 (week 118) and 4;1.4 (week 210). The files presented 1195 questions, marked by the final '?': 654 'real' yes/no questions and 541 content questions (of which 482 were wh-questions with a non-subject wh-constituent). One word utterances, partially intelligible, incomplete or unclear utterances, imitations, immediate repetitions and formulaic routines were left out.

(43) English CP: <+wh> $<_{\text{precedes}}$ <+fin,+aux>



Graph A: Content questions that realize the wh-pronoun: **instantaneous**
Graph B: Content questions for the non-subject that realize <+fin,+aux> in C^o: **one year (50 weeks)**
Acquisition point A → Eureka point B: **18 weeks**

Since child language in English lacks the regular use of the finite Aux in C^o for a considerable time as shown by graph B, leaving the Spec,C underspecified fails to mark the sentence as having a C-projection.[30] Without the functional wh-element, the outcome in child English is a subject-initial clause with a single argument gap, and without C-projection. The child must add the wh-pronoun if there is to be a content question at all. A construction with both the wh-element and the finite Aux absent might have an elliptic effect, but cannot signal a content question.[31] This seems to provoke a much earlier appearance of move <+wh>. The English child meets this challenge and fronts the gap-related wh-element instantaneously, as shown by graph A in (43). Graph A hugs the 100% line from the very beginning.[32] The graphs in (43) show how the acquisition of move <+wh> precedes the acquisition of move <+fin,+aux>

[30] Spec,C stands here for A-bar operator phrase.

[31] Additionally, there is a weird phenomenon to be reported. The earliest content questions are one-word utterances. English toddlers select wh-pronouns for such utterances, e.g. *what* (is this)*?*, *where* (is X)*?*, whereas Dutch children do not. They select a pragmatic sentence adverbial *nou?* ('now?' = 'what is this?', 'where is X?'). See Van Kampen (1997:31,78ff) for the function this sentence adverbial has in child Dutch. It appears in 10% of the adult content questions and in >80% of the child language content questions with wh-drop. Pragmatic adverbials abound in spoken Dutch and German. Children pick them up almost immediately (V. Kampen 2001). Being spoken language, they are hardly analyzed in the standard grammars. For some mysterious reason, they do not appear in English. Cf. the adult Dutch: *nou en?* ('now and?') has as equivalent the adult English *so what?*.

[32] The total percentage of content questions with a wh-element is not completely 100%. This is due to 7 examples of questions, that could possibly be interpreted as content questions, though they lack an auxiliary and a wh-element. See the examples and the reactions of the mother in (i). This type of question covers less than 2% (7/541) of all content questions. They might be no more than slips of the tongue.

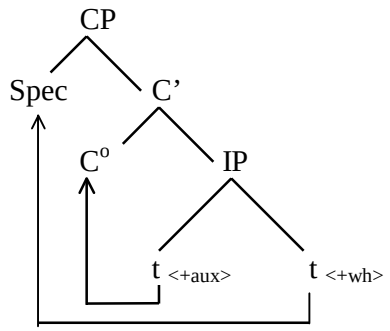
(i) a. *CHI: he saying? (Sarah 2;7.5)
*MOT: he's saying he's a good boy.
b. *CHI: do(se) come from? (Sarah 2;8.2)
*MOT: where'd they come from?

(residual V-second) by almost one and a half year (68 weeks). The next section considers the acquisition of move <+fin,+aux> and the reconstruction of the C^o-I^o chain.

4.2 Relevance of the head chain $C^0_{<+aux>}$ - $I^0_{t<+aux>}$

The acquisition of the C-projection implies that the material in Spec,C and C^o is chain-related to positions within the IP sister constituent of the C^o.

(44)



The recoverability of the empty place t_{wh} follows from the UTAH as in section 3.4. A learning step must relate the empty place t_{wh} to the phrase in Spec,C. The construction of the head-head chain C^o-I^o is not supported by UTAH. Nevertheless, the longitudinal acquisition data give striking evidence that the verbal chain C^o-I^o is perceived from the beginning on.

The input to the acquisition procedure was provided by the mother of English Sarah (CHILDES, Brown corpus). The (arbitrarily chosen) files 11-26 contain input by the mother consisting of 1050 declarative and interrogative sentences with a lexical verb and a subject. The distribution of <+fin,+aux> in this set is given in (45).[33]

(45) *Input percentages* of Aux-distribution in sentences with a lexical verb and a subject
Out of 1050 utterances with a lexical verb and a subject (mother Sarah, Brown corpus)
a. with an Aux in I^o: 45% (479 n)
b. with an Aux in C^o: 25% (257 n)
c. with no Aux at all: 30% (314 n)

Examples (from Sarah 013)
a. her hat doesn't come of
b. what were you doing over at Nana's last night?
c. you sit on it, right

The child's acquisition procedure may induce that <+fin,+aux> elements can be placed in content questions between the wh-phrase and the subject ((45)b), as well between the subject and the predicate ((45)a). The 30% in (45)c seem to indicate that I^o is optional. If <+fin,+aux> is optional in I^o it might be optional in C^o as well. Since most variants in C^o are deeply cliticized variants, they must be hard to spot. If no further insights are derived from the input, all variants in (46) are to be expected in the child's output.

[33] Imperatives were excluded from the count, as well as agreement marking on the lexical verb. Verbal agreement is acquired much later in English child language and unlikely to have been a point of orientation. See also section 4.3.

(46) English child language

Patterns in wh-questions					Attested cases (Sarah, Brown corpus)
a. what	---	you	---	doing?	[child language form]
b. what	---	you	*are*	doing?	[not attested]
c. what	*are*	you	*are*	doing?	[not attested]
d. what	*are*	you	---	doing?	[adult language form]

In fact, English Sarah does start with the double gap pattern (46)a.[34] This follows from the simple tactic of leaving out unidentified functional categories. A year later Sarah ends with the adult target (46)d. Strikingly, the in-between variants (46)b and (46)c make no appearance that has any systematic significance. Variant (46)c does not appear at all and (46)b, the adult form for subordinates, covers only 2% of the material (some 10 attested cases; see also Guasti and Rizzi 1996, Stromswold 1990:tab.5.5-5.6). The same one year acquisition period and the same systematic avoidance of patterns (46)b/c is present in the <+fin,+aux> acquisition of Adam (CHILDES, Brown corpus), as counted by Rowland and Pine (2000). Roughly speaking, Adam's use of pattern (46)b for root wh-questions remains below 10%, although there is one peak measurement of 14.6% (60 cases). However, one comes close to the almost impeccable performance of Sarah under the following reanalysis of the data. Two-thirds of Adam's cases (40 of 60) are brought in by *why* and *how* questions. Children often use these wh-words as variants of *how come*, which also induces subordinate structure. All these subordinates may carry a 'meta-linguistic' intention of clarification. The child asks e.g. '*why* "*we are going to fry some eggs*"?' but what s/he really wants to say 'what does *fry* really mean?' or 'why do you say *fry*?' (Bill Philip p.c., Rowland and Pine 2000:179).[35]

The same chain-oriented acquisition strategy is present in the Aux-movement of Sarah's yes/no questions. We will elaborate on this in section 4.3.

(47) English child language

Patterns in yes/no-questions (*be*-types)				Attested cases (Sarah, Brown corpus)
a. ---	you	---	going home?	[child language form]
b. ---	you	*are*	going home?	[not attested] [36]
c. *are*	you	*are*	going home?	[not attested]
d. *are*	you	---	going home?	[adult language form]

The acquisition procedure is sufficiently informed to avoid the patterns (46)b/c and (47)b/c in root questions.[37] This then is the major problem. How did the acquisition procedure spot that

[34] Guasti and Rizzi (1996) claim that Aux *be*-drop in subject wh-questions would be evaded, whereas it would be regular in non-subject wh-questions for some time. In their view types like *who (is) doing that?* would be rare and types like *what (are) you doing?* would be expected. This conclusion may follow from their theory, but it does not follow from their numbers. They counted (fn.9): 44 non-subject wh-questions with Aux *be* and 303 without (that is, a ratio of 12.7% versus 87.3%); 1 subject wh-question with Aux *be* and 8 without (that is, a ratio of 11% versus 89%).

[35] Indeed, Rowland and Pine (2000:179) report that 12 of Adam's non-inverted *why*-questions 'followed a mother's declarative that modeled the same auxiliary in post-subject position'. The use of non-inverted wh-questions is reported in other studies, e.g. Klima and Bellugi (1967), Cazden (1970). See Ingram (1989:456ff) for an overview. Ingram and Tyack (1979) support the idea advanced here that the use of non-inverted Aux in wh-questions can be put aside.

[36] Valian et al. (1992) report the use of non-inverted yes/no questions, pattern (47)b, as a result of an elicitation task, but no percentages are given. Pattern (47)b is sometimes used in adult English. Sarah's mother uses pattern (47)b with an Aux-in-I in 7% of her yes/no questions. The child Sarah seems to disregard this. We will deal with the other patterns of yes/no questions in (49) and at the end of section 4.3 (fn. 17).

[37] Patterns (46)c and (47)c, with the auxiliary or the (past) tense copied, are reported in the literature by Hurford (1975), Kuczaj (1976) among others. Ingram (1989:459) points out that the nature of the copying facts is

the <+fin,+aux> in C^o heads a chain C^o-I^o where I^o is empty? There seems to be no specific reason to assume an empty, but present, I^o. There are sufficient PF data (30% in (45)c) to suggest that I^o need not be present at all and is an option only. The next question is how the acquisition procedure arrived at the decision that English I^o is present anyway.

4.3 Longitudinal graphs

Functional categories are single morphemes with an extremely high token frequency as compared to lexical content items. Nevertheless, lexical content items may be acquired after a few experiences. The precise semantic and syntactic properties of functional categories (its s-constraints and c-constraints) are not easily identified. They may require tens of thousands of experiences. We propose the functional category enters the reconstruction of the string as the unidentified element (48).

(48) <F?> [c-constraints]
[s-constraints]

The phonological form is not added by the child until the grammatical properties of <F?> have been identified. These are a highly idiosyncratic set of language specific co-occurrence constraints in the predicate. The <F?> has to be allotted a position in the X-bar structure, from where its semantic effects ([s-constraints]) and its syntactic constraints ([c-constraints]) can be seen as some head government effect, like theta/Case assignment was for content items. Since 55% of the input sentences has no I^o at all, see (45)b/c, the acquisition procedure needs a 'bootstrap' to insert a trace I^o automatically.

Fortunately, the *be*-forms, and Aux+negative to the right of the subject are obligatory and indisputable representatives of I^o in the mother's input and in child's output. Suppose, all other representatives of I^o are temporarily disregarded by the acquisition procedure as being inconclusive data for the construction of an obligatory I^o. See (49).

(49)

a. +/– present tense *-s*. Absence of third person *-s* may represent bare stem (root infinitive) or present tense in child English. It is acquired only at the age of six by the child (see Wexler 1994). Hence, it is too unreliable a candidate to orient the acquisition procedure at the age of two and a half.[38]

b. +/– perfect auxiliary *have* (with past participle), +/– past tense. Absence of *have* may represent a <+state> aspectual predicate or past tense in child English. English often carries no distinction in the lexical verb between past and past participle. Hence, they are unlikely to force an obligatory I^o.[39]

c. +/– modal Aux. Bare predicate stems in English child language are either present tense or receive a modal interpretation (see Blom and Krikhaar to appear). Hence, they give no PF indication about the presence of an I^o.

The reduction (49) for I^o representatives leaves us with the *be* and the A*ux-neg* cases. The construction of the longitudinal graphs in (50) is based on them. The graphs show that the <+fin,+aux> in the position I^o (graph D) is identified before the <+fin,+aux> in the position C^o

---

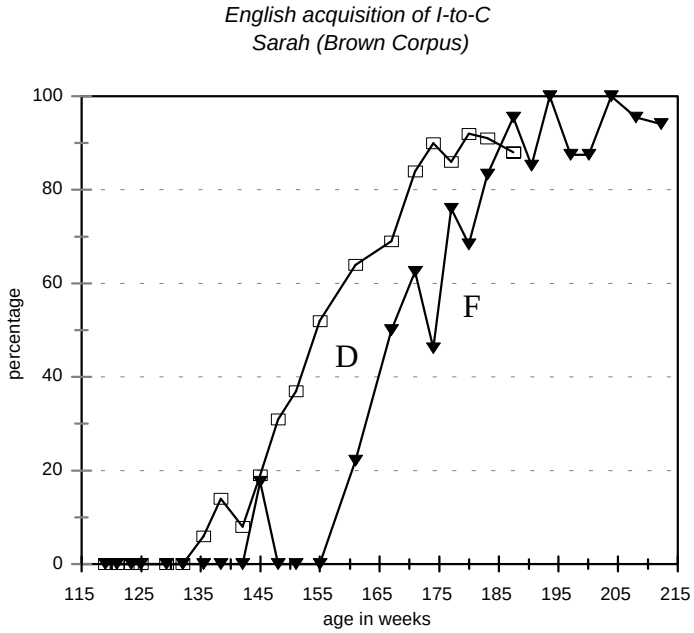
anecdotal. One does not know at what point of development they appear, but they seem to occur quite late. Since we did not find these patterns in the files of Sarah, we will not speculate on it any further.

[38] These I^o forms turn up in the margin of early child language, often in an appropriate context (Borer and Rohrbacher 1997). It needs longitudinal evidence to see whether this morphology is consistently maintained.

[39] For the idea that the past participle/perfect develops from aspectual towards temporal use see for instance Bloom, Lifter & Hafitz (1980), Sano and Hyams (1994), Jordens (to appear).

(graph F). The eureka point and the acquisition point of graph D precede the corresponding points of graph F by some 12-14 weeks.

(50) Obligatory <+fin,+aux> in I^o and C^o



Graph D: Non-questions that realize <+fin,+aux> *be* and Aux-neg in I^o: **31 weeks**
Graph F: Yes/no-questions that realize <+fin,+aux> *be* and Aux-neg in C^o: **29 weeks**
Acquisition point D → Acquisition point F: **12 weeks**

The graphs are constructed from the files of English Sarah.[40] Each of the measuring points that support the acquisition graphs indicates the proportional relation in the files between the target form (obligatory presence of Aux) and its child language counterparts (absence of obligatory Aux). The graphs show a slow, but steady rise. Graph D includes the cliticized forms {'s/'d/'ll/'re}in their various degrees. Together these are frequent.[41] The acquisition graph F reflects the rising use of obligatory <+fin,+aux> in C^o, but under two restrictions. Graph F covers only the C^o elements <+fin,+aux> that: (a) were identified as obligatory I^o representations according to graph D, and (b) were present in yes/no questions. With the first reduction we had in mind that Aux-in-C^o can only be mastered after the acquisition of Aux-in-I^o. The second reduction evades the massive use of cliticized Auxes in the C^o following a wh-pronoun.

The relatively late acquisition of aux-insertion in yes/no questions must be related to 'poverty' of the input. It may be that our approach meets here with the same problems Sarah had to deal with. Sarah's eureka point for yes/no-questions marked by a <+fin,+aux> in C^o, pattern (47)d, trail the eureka point for <+fin,+aux> marking of wh-questions, pattern (46)d,

[40] Sentences without a subject were excluded from the count. Partially intelligible, incomplete or unclear utterances, imitations, immediate repetitions and formulaic routines were also left out. The measurement points are supported by (minimal) 33 and (maximal) 205 cases for graph D; and by (minimal) 11 and (maximal) 50 cases for graph F. Between week 132 and week 187 we counted 1374 cases for graph D and 333 cases for graph F. Graph D then is supported by a fourfold higher amount of data than graph F.

[41] If there had been more material, it would have been better to construct separate acquisition graphs: a graph D1 for non-cliticized *be*-in-I^o and a graph D2 for cliticized *be*-in-I^o. Probably D1 and D2 will remain parallel to the D in (50), one to the left and the other to the right of it. There are additional relevant claims in Brown (1973:306,317) and Stromswold (1990:170). Aux *be* would be acquired after the copula. It would be interesting to see whether there are separate but parallel acquisition graphs for respectively copula *be* and Aux *be*.

by almost half a year (20 weeks). The acquisition point for both <+fin,+aux> markings, the one for content questions (graph B in (43)) and the one for yes/no-questions (graph F in (50)) , reach the acquisition point at the same time (at week 187).[42] The 'poverty' of the yes/no input appears as follows. Some 63% of the yes/no questions brought in by the mother had no subject-aux inversion at all.[43] They were simple statement frames with a question intonation. Preferably, straight *do* was left out, as in *you want a cookie?*[44] So as far as these yes/no questions are concerned, Sarah might have taken the easy way out. In as far as she did, we cannot measure whether she left out a modal or left out a *do*. We can only measure the rise of aux-inversion for yes/no questions by looking at the restricted set of yes/no questions where an auxiliary is obligatorily present and must invert (*be/don't*), as argued for in (49). This development is supported by the 37% of the mother's input that brings out the inverted be-forms, modals and Aux+neg in yes/no questions.[45]

Graph F in (50) is more unruly than the better shaped graph D in (50), but this is hardly remarkable, because D for obligatory <+fin,+aux> in I^o is supported by a fourfold higher amount of data (see footnote 13). Nevertheless, graphs D and F suggest a point of potential interest. The acquisition of <+fin,+aux> in C^o seems to take the same speed as the acquisition of <+fin,+aux> in I^o. If the equal speed of D and F in (50) would be supported by a higher amount of longitudinal data and a few more children, one might consider whether parallel graphs will turn up if a grammatical principle (I^o predicate marking in this case) applies within a new context (root questions in this case). It has been argued by Kroch (1989) that parallel graphs will emerge if a (socio-linguistic) system perceives the possibility of a rule extension (our eureka moment) and starts to adapt itself. Due to the inertia of the social system it may take several generations before it has fully adapted to the new norm (our acquisition point). The adaptation period in language acquisition is measured in weeks, but might operate under a comparable inertia. E.g. the graphs in (40) of Dutch Sarah and Laura (same rules, two

---

[42] This might fit the following picture. All auxiliaries could be used to construct graph B, since in content questions there is only the opposition <+fin,+aux> in C^o or no <+fin,+aux> at all. Suppose there had been sufficient data in graph B to construe separate graphs for each type of <+fin,+aux>. Then it might be the case that all these graphs coincide with and confirm graph B. This result was reached for Dutch wh-movement in Van Kampen (1997:75). Another possibility could be that graph B would split up in a series of graphs that do not coincide, but that are nevertheless all more or less parallel to graphs D and F. Rowland and Pine (2000) have offered some evidence that a graph like B might break up in separate graphs for various sub-cases of the construction. If these graphs were parallel, - and for lack of data nobody knows at the moment -, it would show that there is a frame C^o - I^o and a possibility to move a functional category within that frame from I^o to C^o. Each functional category requires its own 'habituation' period, and all these periods are successive but alike, as reflected in parallel acquisition graphs. For lack of data, this cannot yet be (dis)proved, but it suggests the potential use of longitudinal graphs.

[43] We counted 209 yes/no questions used by the mother of Sarah in the files 11-26, of which only 77 (37%) had an Aux in C^o. Only 'real' yes/no-questions brought in by Sarah and her mother, i.e. questions that asked for an confirmation or a denial, were considered. 'Suggestive' questions, i.e. sentences with a question intonation followed by huh, OK, all right or a precursor of a tag question were left out. These sentences always have the form of a statement.

[44] Namely, 56% of the yes/no questions left out a *do*. A minor percentage (7%) of the uninverted yes/no questions in the speech of Sarah's mother was of pattern (47)b with an uninverted auxiliary, 5% with a cliticized and 2% with a full Aux-in-I, see (i) and (ii).

(i)	your baby's gone home?	(Sarah 12)
	you're not Momma's girl anymore?	(Sarah 14)
(ii)	you don't know?	(Sarah 24)
	you aren't?	(Sarah 24)

The types in (i) and (ii) seem to ask for intimacy rather than information.

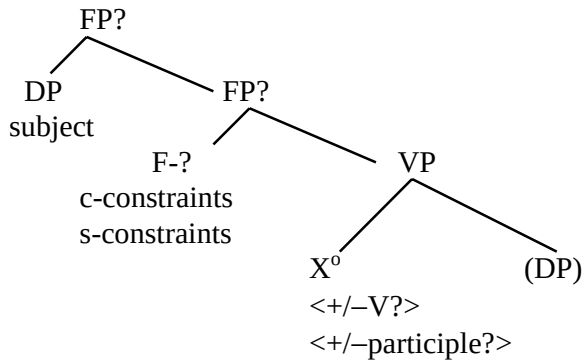
[45] We do not touch the learnability of *do*-constructions here. The early study on the acquisition of *do* by Klima and Bellugi, (1966) needs a remake. It does not offer the data needed for longitudinal graphs.

individuals) are almost parallel and so are the graphs in (50) of English Sarah (one individual, one rule, two contexts). One might think here as well of the inertia of a behavioral system.

4.4 Reconstruction of the verbal chain

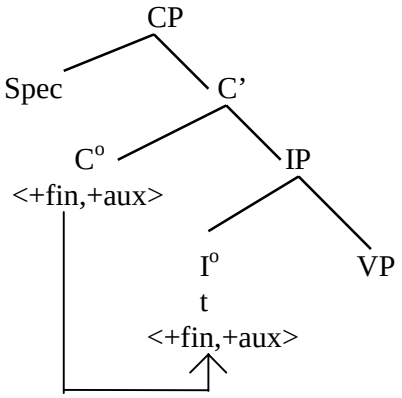
So far, it has been established that the acquisition point of obligatory <+fin,+aux> in I^o precedes the acquisition of <+fin,+aux> in C^o by some 12 weeks. The identification of <+fin,+aux> starts in the position I^o between subject and predicate, see (51). As soon as the highly language specific properties of <F?> have been identified (copula, or present continuous, or passive), its phonologic form can be used with confidence.

(51)



The functional category <+fin,+aux> in C^o starts later. Suppose again that the functional category <+fin,+aux> in C^o starts as a phonologically zero element <F?>. As before, phonologic realization will follow if the c-constraints and the s-constraints of the category have been ascertained. The new semantic constraint must be the illocutive value 'pointed question'. The other syntactic and semantic constraints have already been discovered and must already be present in the category's representation that has just been stored in the lexicon and welded to its phonological form. These are the syntactic category constraints and the semantic constraints that the <+fin,+aux> element imposed on the predicate. They have been acquired with respect to the position I^o. So these I^o configurations may 'come to the mind' naturally. They are induced by the structure preserving conservativeness of the lexicon. This is an easy parallel to Baker's positionally uniform theta assignment UTAH. It is a positionally uniform function application, UFAH so to speak. Since category and aspect constraints on the predicate have already been associated with the position I^o, and since, like UTAH, such pre-established frames can hardly be reversed for any predicate construction, it follows that the category/aspect constraints themselves induce the position I^o between subject and predicate. This position is empty, but can be related to the PF identical <+fin,+aux> element in C^o. The moment is there. The acquisition procedure must insert the $t_{<+aux/+fin>}$ in the position I^o and relate it to C^o. The C^o elements are acquired as chain elements only. See diagram (52).

(52)



Since F [s-constraint,c-constraint] now claims the positions I^o and C^o, the absence of (46)b/c and (47)b/c becomes plausible. Although this seems to us a feasible scenario, it might not have gone unnoticed that its actual execution took a full year and maybe millions of learning experiences. This suggests that such abstract movements as English I-to-C are on the border of learnability.

Further empirical support is the instantaneous appearance of the correct patterns (46)b and (47)b in subordinate questions. The avoidance of (46)/(47)b in root questions combined with its instantaneous appearance in subordinate questions offers an argument that the acquisition of <+fin,+aux> distributions rests on the child's preceding construction of the chain C^o-I^o. It is a general type of argument. All subordinate clause properties are likely to be acquired instantaneously. The argument has already been given for finite verb placement in Dutch subordinates, section 3.4 above. Root clauses are build up from C-movement rules. The C-movement chains imply the underlying structure. Once the acquisition procedure has acquired the root structures, it has acquired the C-movement chains and hence the underlying sentence forms where these chains originate from. The non-root subordinate structures are a simplification towards a pattern that had already been perceived. The chains are rolled up into their foot position. A full competence claim that there is a priori knowledge about Ross' Penthouse Principle (only root structures have C-projections to speak of) could explain the instantaneous acquisition of the subordinate order as well, but by its own logic it would fail to predict the laborious acquisition story of the root structure.

5 Concluding remarks

The C-projections in Dutch and English exemplify the more general acquisition problem for underlying structure. It can be schematized as in (53).

(53) $[[<+wh>]_{Spec}$ $[<+fin>]_{Co}]_{CP}$ […… $t_{<+fin>}$ ….. $t_{<+wh>}$ …..] $_{IP/VP}$

The elements in C relate to gaps in the IP/VP structure. Yet, the acquisition procedure is not supported by information that there have been C-elements extracted at all. Hence it is not clear how a it should establish the true constituency of the IP/VP. The rules that select and reconstruct the data are exactly the piece of grammar that has to be acquired. We have labeled this the Gibson/Wexler paradox. Gibson and Wexler demonstrate how unprejudiced, but otherwise unbelievably well-informed, acquisition procedures can not solve the problem. To that end, they assume a procedure that already commands a full-fledged grammar with all parameters in default setting. The procedure can flip a parameter setting, should that solve a parsing failure of an input sentence. Confronted with input that is to be parsed, the procedure will start holistic jumping and dance up and down in an array of possible core grammars.

Gibson and Wexler subsequently compute that the parameters can always flip in such a way that the acquisition effort runs aground in an incorrigible state. They label such a state as 'local maximum'. In reaction to Gibson and Wexler (1994), others propose strong additional measures in order to escape from such sorry states (see Berwick and Niyogi 1996, Frank and Kapur 1996, Fodor 1998, 2001). For instance, one may assume a procedure that can flip several parameters at the same time. Each sentence will now open a variety of possible analyses and will not be particularly instructive. The procedure keeps track of its more successful combinations and might in the end settle for a final winner without being committed ever to irreversible combinations. Another possibility (Fodor 1998) might be that the procedure is equipped with far more circumspect instructions before flipping a parameter that might lead to a dead end. It is not clear, though, how the procedure gets hold of these safe, but construction specific triggers. Gibson and Wexler (1994) themselves suggest something rather different. They ask neither for an overdose of computational space, nor for result oriented filters. The acquisition procedure, they speculate, might rather suffer from initial limitations in computational space. No more than a temporary failure of neural maturation is what they ask for. This might be sufficient to delay the installment of V-second. In the meantime, the parameters for verb-argument directionality should meet with good fortune and get fixed, irreversibly this time (Gibson and Wexler 1994:433). Some skepticism is justified here as well. Let it be granted that the initial learning steps are bound to be under unusual constraints. It is not clear, though, why the limitation on computational potential should follow from neural 'hardware' rather than from the grammatical 'software'. Further, the Gibson/Wexler modal gives no answer to the general problem of learnability and underlying structure. It offers only a specific example. As far as the specific case is concerned within the model, it remains unclear how lack of neural maturation brings about so specific an effect as the acquisition of the OV directionality before the acquisition of omnipresent V-second distribution. Yet, there is a striking point in the Gibson/Wexler suggestion. The actual acquisition procedure may be more successful than the perceptive linguist by its virtue of being less clever. Less is more!

The actual succession of acquisition steps as demonstrated in longitudinal graphs confirms that, as a matter of fact, the directionality parameters are set first. The reason is not difficult to see if one looks at the data, and it has been seen all along. If an intake filter leaves out the <+fin,+aux> elements as <F?>, the majority of predicate heads with lexical content turn up in the predicate final position. The simple strategy of leaving out as yet unidentified functional categories as <F?>, yields the possibility of setting up a lexicon with content verbs, all associated with an SOV (UTAH) argument frame (sections 3.1.1). The same procedure in English is less surprising, but delivers the SVO (UTAH) frame. The lexicon of content items will now offer a basis to construct a fixed frame IP and VP where argument gaps can be spotted. One might have expected that from this glorious point on, it will be relatively simple to relate the gaps in the IP/VP with the C-elements in front. This does not bear out. Both systems confront the learning procedure with new difficulties. As far as Dutch is concerned, none of the IP/VP root structures indicates the underlying position for <+fin> (section 2). As far as English is concerned, more than half of the IP/VP structures contains no I<+aux> (see (45)b/c, section 4.2). A rash procedure might have decided that there is no empty position in the Dutch IP/VP that is to be related to the <+fin> in C^o. It might also have decided that the I^o <+aux> in English is an option like a sentence adverbial is. These decisions are certainly not taken. The young speaker betrays a keen awareness of the verbal chain C^o - I^o/V^o. Positions that are related by movement are never or hardly ever filled in both. Moreover, the subordinate clauses in both languages shorten the verbal chain and give up the C^o as a verbal position. It is hard to see how the clue of the syntactic root/subordinate distinction could have been grasped without awareness of the verbal chain. As a matter of fact, the acquisition of the subordinate is

instantaneous in both languages (sections 3.4 and 4.4). Yet, this cannot be taken as a strong indication that the verbal chain is imposed on the input string by the a priori conceptions of the learner. There is the simple fact that the acquisition history of the verbal chain in the root clauses is wildly different for both languages. Full V-second takes less than 20 weeks for the Dutch root clause, but the residual V-second for the English root clause takes more than a year. We have tried to offer an input-controlled analysis in both V-second cases (sections 3.2 and 4.2), which implies an input-controlled analysis for the verbal chain.

Besides the reconstruction of an underlying fixed position for I^o in English and Dutch, there is a problem with filling in the C-projection. The C^o head offers a structural bearing for the Spec,C that is to be filled with the <+wh> constituent. The <+wh> constituent appears instantaneously in English child language, while the C^o is still empty or absent. By contrast, the <+wh> constituent appears in Dutch child language after the acquisition of V-second and takes half a year. We have offered a plausibility consideration for this difference. In Dutch child language, the <+wh> element seems initially present as <F?>. The previously acquired finite verb in C^o serves as a reliable identification for the underspecified, Spec,C (section 3.3). Since English child language lacks the overt Aux in C^o for a considerable time, leaving the Spec,C underspecified as <F?> fails to mark the sentence as having a C-projection (section 4.1.1). It seems to confirm the earlier impression from the V-second phenomena that the acquisition procedure is highly dependent on the twists and turns of the input.

From a less construction specific point of view, one may see the development of child language as a progression through the series of grammars G, all derived from and provoked by the same input.

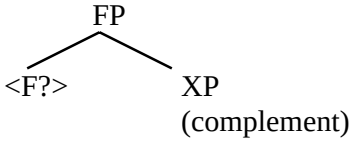
(54) G: G_o .., G_{i-1}, G_i, .. G_n

The linear succession in G can be factually demonstrated by the linear order of longitudinal graphs for the acquisition of the various functional categories, e.g. <+fin>, <+wh>, <+aux>. The linear order in G will follow from the collaboration of the two factors in (55)

(55) a. an I-language factor: the hierarchy of grammatical categories <+F>
b. an E-language factor: the identifiability of the <+F> complement

The two factors can be seen as feeding a learning function f_i , that will produce the functional category F_i under appropriate circumstances: f_i (x) = F_i. What f_i needs are (a) the availability of G_{i-1} for the I-language factor, and (b) an input data set D_i for the E-language factor. The data set D_i should consist of structures of type (56).

(56)



The complement XP must be analyzable by G_{i-1} in some 70% of the <F?> cases. As soon as <F?> has been identified as F_i , the grammar G_{i-1} and its new addition F_i are renamed G_i. A learning function f_i elaborated with these ingredients looks as: f_i (G_{i-1}, D_i) = F_i. The learning function f_i operates within the wide domain of UG possibilities. It must identify the s-constraints and the c-constraints that F_i will impose on the structure it applies to. That is, it must find a systematic cognitive point of view (the s-constraint) and a systematic distributional point of view (the c-constraint). These two defining content properties of F_i derive from the complement XP as analyzed by G_{i-1}, i.e. they are input-controlled. The

linearity of the acquisition series G follows as well from properties of the grammatical system. The order in G derives from the asymmetric head-complement application hierarchy of functional categories, i.e. the learnability hierarchy G follows from a property that is inherent to the system G_n that is being confronted. The factor G_{i-1} functions as an evidence filter. It enforces a systematic but benign neglect of all functional categories that have a downstairs context that is not analyzable by the current grammar G_{i-1}. Higher level F_i/FP cannot be perceived and identified with sufficient clarity by G_i, unless its lower level complement XP has been analyzed first by G_{i-1}. The acquisition procedure is forced to climb up the grammatical hierarchy embodied in the grammar it confronts. This is not a particularly sophisticated arrangement. The factor G_{i-1} within f_i simply defines the current amount of ignorance about G_n. So far about grammatical hierarchy, the I-language factor in (55).

As has been demonstrated by the C-projections in Dutch and Engliah, it is by no means guaranteed that an initial neglect of higher functional categories like <+fin>, and <+wh> cannot fail to deliver a remnant structure XP with suitable data to set the major parameters. This is totally dependant on the E-language factor mentioned in (55). E-language makes that the evidence filters from grammatical hierarchy just happen to work. The E-language proportions happen to solve the Gibson/Wexler paradox, not in principle but in practice. The present acquisition strategy inevitably implies that V-movements, verbal movement chains, cannot be learned unless the I-language contains the overt category <+aux> and the E-language maintains its use for 70% of the predications. This looks like a bad omen for abstract syntax. Another property of Chomskyan grammar is captured much better. Abstract grammatical functions for argument structure, predication and scope assignment are represented as category hierarchies in phrase structure and learned in that order. This is in fact what the linear order in G (54) is building up. Finally, we repeat the three conclusions that have been promised at the end of section 1.

(i) *Facts*
There is a parameter setting order. The order and the relative speed of parameter setting can be demonstrated by linear acquisition graphs of the type introduced by Brown (1973). The actual order of parameter setting is as expected by Gibson and Wexler (1994). Paradoxically, underlying directionality comes first.

(ii) *Acquisition procedure*
The acquisition order and the acquisition speed of move <+wh/C> and move <+fin/C> in Dutch and English are different but due to the same evidence filters on the input. There are quantitative filters that depend on frequencies in E-language. They allow an uninformed language acquisition procedure to perceive underlying constructs first and transformationally derived constructs later. The incremental stages follow from I-language hierarchies highlighted by E-language frequencies.

(iii) *Metaphysics*
The learnability of natural grammar follows from an interaction between I-language and E-language. Both factors are imposed on the learner. The present proposal stresses control by input and shies away from Chomskyan mentalism where wired-in properties of the human brain itself lay down their a priori conceptions on a cultural product. In the present view, it is different. Combinatorial systems are inventions by human culture imposed on a (tabula rasa) brain. The UG system and its natural grammars are natural in the sense of ‘natural numbers’. Objective mind defies grey matter.

References

Baker, M. (1988) *Incorporation. A Theory of Grammatical Function Changing*, Chicago: University of Chicago Press.

Behrens, H. (1993) *Temporal Reference in German Child language. Form and Function of Early Verb Use*, PhD. dissertation University of Amsterdam.

Berwick, R. and P. Niyogi (1996) 'Learning from triggers', *Linguistic Inquiry* 27.4, 605-622.

Besten, H. den (1977/1983) 'On the interaction of root transformations and lexical deletive rules', in: E. Abraham (ed.) *On the Formal Syntax of Westgermania*, 47-131, Amsterdam: Benjamins.

Blom, E. (2002) *The Development of Tense and Agreement in Dutch Child Language*, PhD dissertation Utrecht University.

Blom, E. and E, Krikhaar (to appear) 'Where to use a nonfinite clause? An experiment on the use of verb forms in early Dutch and English child language', to appear in: *Proceedings of GALA* 2001.

Bloom, L., K. Lifter and J. Hafitz (1980) The semantics of verbs and the development of verb inflections in child language', *Language* 56, 386-412.

Borer, H. and B. Rohrbacher (1997) 'Features and projections: Arguments for the full competence hypothesis', in: E. Hughes, M. Hughes and A. Greenhill (eds.) *BUCLD 21 Proceedings*, 24-35, Somerville MA: Cascadilla Press.

Braine, M.D.S. (1963) 'The ontogeny of English phrase structure: the first phase', *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.

Brown, R (1973) *A First Language: the Early Stages*, Cambridge Mass.: Harvard University Press.

Cazden, C. (1970) 'Children's questions: their form, their function and roles in education', *Young Children*, March, 202-220.

Chomsky, N. (1981) *Lectures on Government and Binding*, Dordrecht: Foris.

Chomsky, N. (2000) *New Horizons in the Study of Language and Mind*, Cambridge Mass.: Cambridge University Press

Clahsen, H. (1982) *Spracherwerb in der Kindheit*, Tübingen: Narr.

Clahsen, H. (1991) 'Constraints on parameter setting. A grammatical analysis of some acquisition stages in German child language', *Language Acquisition* 1.4, 361-391.

Clark, R. and I. Roberts (1993) 'A computational modal of language learnability and language change', *Linguistic Inquiry* 24-2, 299-345.

Deprez, V. and A. Pierce (1994) 'Crosslinguistic evidence for functional projections in early child language', in: T. Hoekstra and B. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar*, Amsterdam: John Benjamins.

Fodor, J.D. (1998) 'Unambiguous triggers', *Linguistic Inquiry* 29.1, 1-36.

Fodor, J.D. (2001) 'Setting syntactic parameters', in: M. Baltin & C. Collins (eds.) *Handbook of Contemporary Syntactic Theory*, 730-767, Oxford: Blackwell Publishers.

Frank, R. and S. Kapur (1996) 'On the use of triggers in parameter setting', *Linguistic Inquiry* 27.4, 623-660.

Gibson, E. and K. Wexler (1994) 'Triggers', *Linguistic Inquiry* 25.3, 407-454.

Gleitman, L. (1990) 'The structural source of verb meaning', *Language Acquisition* 1, 3-55.
Grimshaw, J. (1993) 'Minimal projection, heads, and optimality', *Technical Report* 4, Rutgers University center for Cognitive Science.
Guasti, M. and L. Rizzi (1996) 'Null Aux and the acquisition of residual V2', in: A. Springfellow, et al. (eds.) *BUCLD 20 Proceedings*, 284-296, Somerville MA: Cascadilla Press.
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.
Hurford, J. (1975) 'A child and the English question formation rule' *Journal of Child Language* 2, 299-301.
Ingram, D. (1989) *First Language Acquisition, Method, Description and Explanation*, Cambridge: University Press.
Ingram, D. and D. Tyack (1979) 'The inversion of subject NP and AUX in children's questions', *Journal of Psycholinguistic Research* 4, 333-341.
Jakobson, R. (1942) *Kindersprache, Afasie und allgemeine Lautgesetze*, Uppsala Universitets Årsskrift 9, Uppsala: Uppsala Universitet.
Jordens, P. (to appear) 'Finiteness in early child Dutch', to appear in *Linguistics*, September 2002.
Kampen, J. van (1992) 'Underspecification of functional features. The acquisition of root wh-questions in Dutch and English', paper presented at the 'Boston University Conference on Language Development', MIT, October 23-27.
Kampen, J. van (1997) *First Steps in Wh-movement*, Delft: Eburon.
Kampen, J. van (2001) 'Pragmatic bootstraps for category assignment: from non-discourse towards discourse oriented language', in: *Proceedings of ELA* (Early Lexicon Acquisition), cd-rom.
Kampen, J. van and F. Wijnen (2000) 'Grammaticale ontwikkeling', in: A.M. Schaerlaekens and S. Gillis (eds.) *Kindertaalverwerving*, 225-285, Groningen: Wolters-Noordhoff.
Kayne, R.S. (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge Mass.
Klima, E. and U. Bellugi (1966) 'Syntactic regularities in the speech of children', in: J. Lyons and R.J. Wales *Psycholinguistic Papers*, Edinburgh: University Press, 183-208.
Kroch, A. (1989) 'Reflexes of grammar patterns of language change', *Journal of Language Variation and Change* 1.3, 199-244.
Kuczaj, S.A. (1976) 'Arguments against Hurford's Aux copying rule', *Journal of Child Language* 3, 423-427.
Lebeaux, D. (1988) *Language Acquisition and the Form of Grammar*, PhD. dissertation University of Massachusetts.
Levin, B. and M. Rappaport (1986) 'The formation of adjectival passives', *Linguistic Inquiry* 17-4, 623-661.
Lightfoot, D. (1991) *How to Set Parameters: Arguments from Language Change*, Cambridge Mass.: MIT Press,.
Lightfoot, D. (1999) *The Development of Language: Acquisition, Change and Evolution*, Malden MA: Blackwell.
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.

MacWhinney, B. (1991) *The CHILDES Project: Tools for Analyzing Talk*, Hillsdale, New York: Lawrence Erlbuam.

Nespor, M., M.T. Guasti and A. Christophe (1996) 'Selecting word order: the Rhythmic Activation Principle', in: U. Kleinhenz (ed.) *Interfaces in Phonology*, Berlin: Akademie Verlag.

Penner, Z. (1993) 'The Earliest stage in the acquisition of the nominal phrase in Bernese Swiss German', *Arbeitspapiere Universität Bern* 30, Bern: Institut für Sprachwissenschaft.

Penner, Z. (1994) 'Asking questions without CPs? On the acquisition of root wh-questions in Bernese Swiss German and standard German', in: T. Hoekstra and B. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar*, 177-213, Amsterdam/Philadelphia: Benjamins.

Poeppel, D. and K. Wexler (1993) 'The full competence hypothesis of clause structure in early German', *Language* 69.1, 1-33.

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, 37-77.

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: Empirical Findings, Theoretical Considerations and Cross-linguistic Comparisons*, Amsterdam: Benjamins

Roeper, T. (1972) *Approaches to a Theory of Language Acquisition with Examples from German Children*, PhD. dissertation Harvard University, Cambridge Mass.

Roeper, T. and J. Weissenborn (1990) 'How to make parameters work', in: L. Frazier and J. de Villiers (eds.), *Language Processing and Language Acquisition*, 147-162. Dordrecht: Kluwer.

Rohrbacher, B. and A. Vainikka (1995) 'Verbs and subjects before age 2: The earliest stages in Germanic L1 acquisition', in: J.N. Beckman (ed.) *Proceedings of NELS* 25, 55-69, Amherst MA: GLSA.

Rowland, C and J. Pine (2000) 'Subject-auxiliary inversion errors and wh-question acquisition: what children do know?' *Journal of Child Language 27(1)*, 157-181.

Ruhland, R. (1998) *Going the Distance. A Non-linear Approach to Change in Language Development*, PhD. dissertation Groningen University.

Ruhland, R., F. Wijnen and P. van Geert (1995) 'An exploration into the application of dynamic systems modeling to language acquisition', in: *Amsterdam Series in Child Language Development* 4 'Approaches to parameter setting', 107-134, Amsterdam: Department of Linguistics University of Amsterdam.

Sano, T. and N. Hyams (1994) 'Agreement, finiteness and the development of null arguments', in: *Proceedings of NELS* 24, 543-558.

Santelmann, L. (1995) *The Acquisition of Verb Second Grammar in Child Swedish: Continuity of Universal Grammar in Wh-questions, Topicalization and Verb Raising*, PdD. dissertation Cornell University.

Saussure, F. de (1915) *Cours de Linguistique Générale*, publié par Ch. Bally, A. Sechehaye and A. Riedlinger, Paris: Payot (1964).

Schlichting, L. (1996) *Discovering Syntax*, PhD. dissertation Katholieke Universiteit Nijmegen.

Stromswold, K. (1990) *Learnability and the Acquisition of Auxiliaries*, PhD. dissertation MIT.

Tomasello, M. (1999) Paper presented at the symposium 'Theory and data: the distribution of forms in arguments', VIIIth International Congress for the Study of Child Language, 12-16 July.

Tracy, R. (1994) `Raising questions: Formal and functional aspects of the acquisition of wh-questions in German', in: E. Lattey and R. Tracy (eds.) *How Tolerant is Universal Grammar?*, 1-34, Tübingen: Niemeyer.

Valian, V. (1990) 'Null subjects: a problem for parameter-setting modals of language acquisition', *Cognition* 35, 105-122.

Valian, V., I. Lasser and D. Mandelnaum (1992) 'Competing analyses of children's early questions', paper presented at the 17th Boston University conference on Language Development, Boston Mass., October 23.

Watanabe, A. (1993) *AGR-based Case Theory and its Interaction with the A-bar-system*, PhD. dissertation MIT.

Watanabe, A. (1994) 'The role of triggers in the extended split Infl hypothesis: Unlearnable parameter settings.' *Studia Linguistica* 48, 156-178.

Weerman, F. (1989) *The V-second Conspiracy*, PhD. dissertation Utrecht University.

Weinberg, A. (1990) 'Markedness versus Maturation: The Case of Subject-Auxiliary Inversion', *Language Acquisition* 1.2, 165-194.

Weissenborn, J. (1991) 'Functional categories and verb movement: The acquisition of German syntax reconsidered', in: Rothweiler (ed.) *Linguistische Berichte* Sonderheft 3 `Spracherwerb und Grammatik', 190-224, Göttingen: Westdeutscher Verlag.

Wexler, K. (1994) `Optional infinitives, head movement and the economy of derivations', in: D. Lightfoot and N. Hornstein (eds.) *Verb Movement*, 305-350, Cambridge Mass.: Cambridge University Press.

Wexler, K. (1998) 'Methodology in the study of language acquisition: A modular approach.', in: W.C. Ritchie and T.K. Bhatia (eds.) *Handbook of Language Acquisition*. Elsevier Academic Press.

Wijnen, F. (1995) Incremental acquisition of phrase structure', in: J.N. Beckman (ed.) *Proceedings of NELS* 25 (vol. 2), Amherst Mass: GLSA, 105-120.

Wijnen, F. (1997) 'Functionele categorieën in Nederlandse kindertaal.' *Nederlandse Taalkunde* 2, 178-198.

Wijnen, F. (1998) 'The temporal interpretation of Dutch children's root infinitivals. The effect of eventivity', *First Language* 18, 379-402.

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: Centre for Language and Cognition.

Zwart, C. J.-W. 1996. *Morphosyntax of Verb Movement: a Minimalist Approach to the Syntax of Dutch*. Dordrecht: Kluwer.