Review of Becker 'The Development of the Copula in Child English'
by *Jacqueline van Kampen*

1.   Theoretical assumptions

Suppose that first language acquisition were like filling in options on an a priori given frame called Universal Grammar (UG). The UG frame is to fit all languages. For that reason UG will open more options than the specific grammar that is being learned. Fortunately, no matter how elaborate the frame is, the child will eventually find and make the restrictive options that fit the primary language data it is confronted with. The more elaborate orientation by UG might nevertheless betray itself. The child may temporarily apply grammatical distinctions that are provided by UG, although it turns out later on that such distinctions are glossed over by the particular grammar that is being learned. Child grammar, of course, is notorious for its temporary underspecifications, rather than for its temporary overspecifications. Nevertheless, overspecifiations cannot be due to input control and for that reason constitute the better arguments for the UG control of the acquisition. Becker's thesis proposes that such an overspecification can be found in the primary acquisition of the English copula. Some copula constructions are more inviting for the child to insert the copula in, although the input itself is impartial and recommends, so to speak, the copula for all non-verbal root predicates. The child's preferential choices are, Becker claims, not accidental, but betray a distinction glossed over in the English PF, but clearly a UG distinction if we look at the findings of comparative grammar.

Becker's thesis alternates between the two perspectives, UG and comparative grammar versus language acquisition. Firstly, following the relevant literature, Becker argues that the UG of non-verbal predicates is sensitive to the semantic distinction between properties intended to be permanent/inherent (as *is a black horse*) and properties intended to be temporary/non-inherent (as *is in the stable*). If the lexical predicate head in copular constructions assigns to the subject a permanent/inherent quality it is said to be an 'individual-level' predicate, if it assigns a temporary quality it is said to be a 'stage-level' predicate (Carlson 1977). The same distinction can be drawn within verbal predicates: eventive verbs are stage-level and stative verbs are individual-level. The stage-level predicate is compatible with all kinds of additional aspectual qualifications, like *is -ing / for a while*. The individual-level predicate does not allow aspectual qualifications, or does so under very restrictive conditions only. That is, the predicate head is marked either <+Asp> as in (1)a, or <−Asp> as in (1)b.

(1)
a.   Dolly is [on the table]$_{<+Asp>}$
b.   Beauty is [a black horse]$_{<−Asp>}$

The stage-/individual-level distinction is relevant for several syntactic environments. It plays a part in small clause complements of perception verbs, in existentials, in *when*-clauses. There are various proposals to account for the syntactic differences (Diesing 1992, Kratzer 1995, among others). Elaborating a proposal by Heycock (1995), Becker argues that the <+/− Asp> property of predicates is best captured by a distinction in phrase structure. Stage-level predicates would project an Aspect Phrase and individual-level predicates would not (p.106-113).

The <+/−Asp> of predicates may become relevant for subject licensing in language typology. From an extensive comparison of various copula constructions in Russian, Hebrew, Spanish, Portuguese and African American English, Becker draws the conclusion that the

<−Asp> individual-level predicates require explicit subject licensing by a <+tense> element, such as the copula. By contrast, the copula may stay out for <+Asp> stage-level predicates. The semantic <+Asp> context of the predicate would already offer a temporal orientation for the sentence (Hebrew). Alternatively, the <+Asp> non-verbal predicate may rely on other solutions: a specific copula (Spanish, Portuguese *estar*) or an instrumental Case (Russian). The various languages draw the line between permanent and temporary properties in a slightly different way according to the syntactic devices they can deploy, but the major distinction between individual-level and stage-level predicates remains (p.176).

Becker's conclusion about UG orientation in copula selection sets the stage to enter the second perspective: primary copula acquisition in standard English. Becker proposes that the universal distinction stage-level/individual-level is reflected in the different acquisition speeds of the copula for English. The copula for <+Asp> predicates is acquired later than the copula for <−Asp> predicates. Becker counts the non-verbal root predicates in the files of four children and this yields a set of 1.597 predicates that would require copula insertion for the adult target language (p.134). In the beginning both <+Asp> and <−Asp> predicates will lack the copula. In the end, both type of predicates will support a copula. Somewhere in the middle, between those two moments, there must be the evidence that the <−Asp> predicates have taken the lead in copula insertion. Now Becker makes a striking move. Rather than tracking down the longitudinal development for each of the children (for which she has counted the number of cases per file) over the crucial period, she simply divides the total 1.597 predicates in <+Asp> and <−Asp> and compares the percentages for copula insertion. It turns out that some 71 % of the <−Asp> have the copula, versus some 34 % of the <+Asp> predicates. This is only possible if the <+Asp> seriously lag behind in copula insertion as compared with the <−Asp> predicates, as Becker's UG thesis anticipated.

Becker's thesis is a well-written and highly instructive piece of work. It contains an excellent exposition on stage-level and individual-level predicates. It gives an interesting modeling of this distinction in present day abstract syntax. It makes a serious attempt to find comparative evidence for a UG conjecture and, above all, it gives an excellent quantification of copula acquisition facts in English. However, I am not quite convinced by Becker's evidence and I will show why. Firstly, I will discuss the method used by Becker to analyze the acquisition data of the children, and subsequently the typological conclusions she draws from the adult languages.

2.   Evidence from acquisition: Nominal and locative predicates

In English, the stage-level/individual-level split roughly divides nominal <−Asp> predicates from locative <+Asp> predicates. All nominal predicates are individual-level and all locative predicates are stage-level (p.62). Becker now claims that in English child language these two types of predicate differ in the acquisition speed of their copula. This difference in acquisition speed would reflect a difference in UG phrase structure requirements (p.97). I will look more in detail which acquisition facts support this conclusion. The acquisition data of the child Nina (Suppes corpus), who displayed the largest contrast between nominal and locative predicates, will be taken as an example.

2.1.   Locative predicates

Becker uses the following method. She considers positive facts of overt *be*, for which she counts the percentage per file and then she averages the percentages across files (p.89 fn.10). Becker correctly states that the average rate of overt *be* across several files "takes into account potential changes in the development over time".

Elegant though this approach may be, it leaves two points in the dark. Firstly, the argument from averages of percentages does not depict the actual speed of the two types of copula insertions. Secondly, the method does not consider how such differences as there are, might have been caused by input factors besides or instead of UG preferences. The relevance of the first point will be shown in a moment by longitudinal graphs of Nina for the locative predicates. The second point is more obvious in a direct way. If one considers the adult target, i.e. Nina's mother, it appears that overt copula constructions are not the only relevant forms. The target itself contains numerous locative <+Asp> predicates without copula, for instance a) in 'small clause' complements with perception verbs, b) in 'small clause' complements with eventive verbs (e.g. *put*, *want*, *take*, *try*, *hold*, *lie*, *get*, *leave*, *keep*), c) in existentials, and d) in constructions with a locative verb (e.g. *fit*, *come*, *go*, *lie*, *sit*, *stand*, *live*, *sleep*). See (2).

(2)
a. did you see *the mouse on the TV* ?   (mother (Nina 04))
b. why don't you put *the frog on the doggie*?   (mother (Nina 09))
c. there're *teddy bears on the box*   (mother (Nina 07))
d. *Janet* went *on the swing* too?   (mother (Nina 12))

Nina reacts to this input with 'bare' predications as in (3).

(3)
a. mouse on TV   (reaction Nina)
b. frog on the sweater   (reaction Nina)
c. teddy bears on it   (reaction Nina)
d. Janet on the swing   (reaction Nina)

As a matter of fact Nina is overflowed with input containing <+Asp> locative predicates in adult constructions like in (2)a,b. Moreover, small clause subject and small clause predicate probably carry the important marks of sentential stress and by far outshine the grammatical context that licenses them in the adult (and UG) grammar (Van Kampen 1997, chapter 2). Nina is bent on picking out the most relevant points by means of two phrase utterances [subject$_{content}$-predicate$_{content}$]. Nina's copula reductions might have been inspired by UG orientation, but the input itself must have had some direct impact as well. If it were a primary concern of Nina to get fully grammatical sentences within a limited frame, she might have used a (zero) copula construction. However, this would often not convey Nina's intention. E.g. *another button on the snowman* (Nina 11), is not at all meant as 'another button is on the snowman'. As a matter of fact the button is still a piece of clay lying on the table.

The true question is whether we can depict the acquisition of the licensing context of (non-verbal) predicates, copulas as well as other contexts. This brings me to the first point mentioned above, the actual acquisition speeds. The longitudinal graph B of Nina in (4) depicts the rise of licensing context for the <+Asp> predicates in (2)a,b (small clause complements of a full verb) and graph C the rise of copula insertion. Together they are counterbalanced by the decline of 'bare' locative predicates, graph A. Graph A is simply the reverse of graphs B+C.

(4)

<+Asp> locative predicates
Nina (Suppes Corpus)
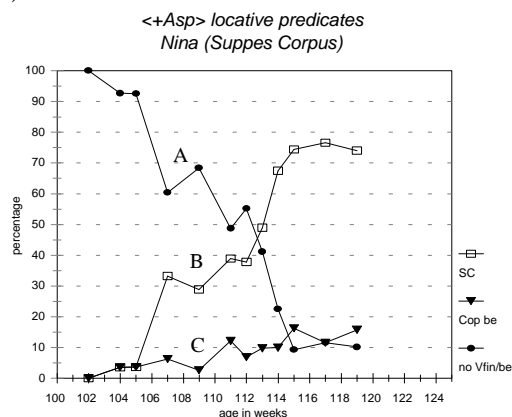
Number of examples: 22 - 55 - 27 - 57 - 38 - 41 - 29 - 51 - 40 - 43 - 43 - 38

Graph A: locative predicates that do not realize the finite verb: 'doggy on the table'
Graph B: locative predicates in 'small clause' constructions: '*I put* the doggy on the table'
Graph C: locative predicates that realize the copula *be*: 'the doggy *is* on the table'

Becker analyzed the part in the middle, week 107-114, the period of optionality. My longitudinal graphs start before and end after that period. Existentials and locative verb constructions ((2)c,d) were left out. The licensing context of existentials appear predominantly as bare *there's X(P)* in early child language, clearly a fixed phrase. Nina produced only 4 existentials with a locative in the coda. The licensing context by locative verbs constitutes a rather complicated acquisition issue. If locative verb contexts would have been taken into account, it would have further enhanced the rising speed of graph B and would have diminished the rise of C, but without a change in the overall picture. The overall picture in (4) is that the insertion of the copula (graph C) plays a minor role in the licensing of locative predicates, since locative predicates predominantly become 'small clause' complements (graph B). Becker was certainly aware of the potential problem of B and C, but because her attention was directed at copula insertion, rather than at predicate licensing in general, and because she did not construct longitudinal graphs, but rather trusted the elegant argument of the proportional differences between predicates with/without copula, she underestimated the quantitative weight of the problem.

2.2. Nominal predicates

There are two striking peculiarities for the acquisition of nominal predicates, both mentioned by Becker. Firstly, nominal predicates mainly occur with pronominal subjects, whereas locative predicates appear in almost half of the cases with full DP subjects/proper nouns (p.147). Secondly, nominal predicates appear in early child language largely as deictic presentationals that give a name to a subject: *that's a horse* (p.151). Presentational statements are obviously the most frequent type of nominal predicates in child language.

If one looks at <−Asp> nominal predicates during the entire learning period of Nina (week 102-115) the following picture arises. Nina's nominal predicates had for 100% pronominal subjects, of which 96% was of the presentational pointing type, predominantly *that('s) X* (84%). The remaining 4%, i.e. with other pronoun subjects, like *he('s)/it('s)/I (am)*, is far beneath a reliable measuring potential. The fixed formula *that's X* (60%) already in itself accounts for the high percentage of *-s* in the <−Asp> predicates. For instance, the first file of Nina contained the nominal predicates in (5) only.

(5) Patterns and numbers of nominal predicates in Nina 01 (week 102)
a. *that's X*: 7
b. *this X*: 6
c. *this is X / those are X*: 1 / 1

The copula is clearly present in the patterns in (5)c, clearly absent in (5)b. The presence in pattern (5)a is a matter of taste. The *-s* clitic in *that's X* can easily occur as an analogue to the *-s* of *this* (and *those/these*). The form *thats* is now seen an unanalyzed whole. Of course this observation is not new (Brown 1973, 307). Becker (p.166, note 5) admits this possibility for the existential pattern *there's X*, but does not consider it for the *that's X* presentationals.

Now compare the longitudinal graphs A and B in (6). Graph A in (6) shows the disappearance of the <–fin> locative predicates, the <+Asp> ones: *mouse on TV → the mouse is on TV/I put the mouse on TV* (cf. graph A in (4)). Graph B shows the disappearance of the <–be> nominal predicates, the <–Asp> ones: *that a black horse → that's a black horse*.

(6)
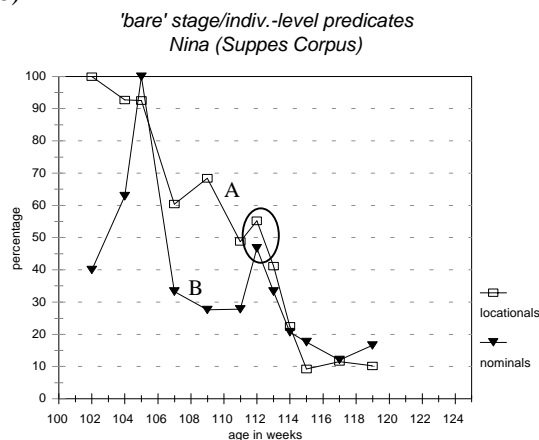
Number of examples graph A: 22 - 55 - 27 - 57 - 38 - 41 - 29 - 51 - 40 - 43 - 43 - 38
Number of examples graph B: 15 - 29 - 17 - 28 - 29 - 18 - 15 - 24 - 29 - 41 - 33 - 48

Graph A: no licensing context in <+Asp> locative predicates
Graph B: no licensing context in <–Asp> nominal predicates

I suggest the following interpretation. The two graphs more or less coincide from the encircled point at week 112 on. Before that point graph B is quite unruly. This might be due to the fact that *thats* has not yet been identified as a contraction of *that+is*. The form *thats* is still seen as a fixed formula for the presentational type. When the copula appears more or less regular in the other copula constructions, *thats* becomes analyzable as *that+is*. The copula in *that+is* then joins at week 112 the general development in A that might be characterized as licensing non-verbal predicates by a <+V> element, a copula or another small clause verb. And thereafter the acquisition of *be* follows as a uniform development. One wonders whether longitudinal analyses of other children might show a parallel picture.

3.   Adjectival and verbal predicates

Becker turns next to the acquisition of verbal and adjectival predicates in English. In contrast to nominal and locative predicates, the individual-level/stage-level split runs within the set of adjectival and within the set of verbal predicates.

3.1. Adjectival predicates

Adjectival predicates can either be <+Asp> stage-level or <−Asp> individual-level. If the same discrepancies would arise within the acquisition of adjectival predicates, the UG control of copula insertion would gain in plausibility. Again, Becker counts the percentage of overt *be* per file and then averages the percentages across files. And indeed, the different rates of *be* acquisition for both types were more or less confirming Becker's proposal (p.135). There was a greater tendency for an overt copula to occur with individual-level than with stage-level adjectives (68% versus 46%). However, the percentages Becker used to come to her conclusions for adjectival predicates were often build upon less than 10 cases per file, in my opinion an insufficient amount of cases. One might demand of this method percentages of samples of some size, cf. the numbers for the graphs in (4) and (6). This is not nit-picking. The number of cases varied greatly between files, between 1 and 17 for the individual-level adjectives, between 3 and 23 for the stage-level adjectives (p.223). This makes the averages across files unreliable to my mind. If there would have been more examples, say 20 to 50 per file, one could use the longitudinal method in order to construct graphs comparable to the ones in (4) and (6). However, as things are now, there is no sufficient basis for quantitative arguments.

3.2. Verbal predicates

The same semantic distinction <+Asp>/<−Asp> extends to verbal eventive predicates (stage-level) and verbal stative predicates (individual-level) (Kratzer 1995). Becker now addresses the issue whether English child language extends the <+tense> requirement on individual-level predicates to stative verbs (p.142). In the literature on Dutch and German child language it has indeed been argued that eventive verbs tend to be nonfinite, whereas stative verbs tend to be finite (Jordens 1990). Becker concludes that the prediction is not met for English child language, but I find this hardly surprising. Firstly, finiteness on the lexical verb in English is a tricky and problematic question. It is too unreliable a candidate in the input to orient the child at a young age. Since positional evidence is lacking, absence of third person *-s* may represent bare stem (root infinitive) or present tense in child English. The percentages of finiteness on lexical verbs can be read off from Becker's tables (p.225). None of the four children showed a clear rise in development of third person *-s*. The percentages of Nina are particularly revealing in this respect. Nina showed the following anti-development (for percentages based on over 30 examples) from week 107 till week 114: 31%→13%→8%→6%→5%. Secondly, Becker does not generalize over all verbal predicates, but treats progressive predicates separately. Progressives are always stage-level eventive predicates. Bare progressives in child English are perfectly comparable to bare infinitives in child Dutch, compare the English input 'is *walking*' with the Dutch input 'is aan het *lopen*'. So they are relevant for a comparison between the licensing of <+Asp> and <−Asp> predicates.

There is another point to consider. It concerns the adult target. Adult Dutch makes the same distinction between stative and eventive verbs in the verbal distribution in finite root sentences. Stative verbs appear by and large in finite V-second position. Eventive verbs are used abundantly with an auxiliary/modal and appear in nonfinite final position (Schlichting 1996). Here again, the input itself might have had some direct impact as well.

4.   Cross-linguistic evidence

Becker's conjecture that <+Asp> predicates do not need the copula is best illustrated by Hebrew. Strictly speaking Spanish/Portuguese, Russian, and African American English do not have to offer much support. Spanish and Portuguese use two copulas, *ser* and *estar*, to encode a syntactic distinction between stage-level and individual-level non-verbal predicates (p.63), and Russian uses two Case-markings on the predicate, nominative and instrumental (p.183). African American English uses a non-finite *be* in specific <+Asp> habitual contexts. The effects of zero encoding of stage-level predicates are rather weak in that language (p.203).

In Hebrew present tense predicates, the zero marking competes with a real copula, which has the form of a 3rd person pronoun (p.71). All non-verbal predicates allow this 'copula' to be omitted. Although, the situation is quite complex for Hebrew, globally the following situation arises. The 'copula' is obligatorily present in predicates with a generic/inherent reading and absent (optional or obligatory) in predicates with a non-generic/non-inherent reading (p.175).

In view of this limited basis in comparative grammar (Hebrew only), it is unfortunate that Becker has not come across Stassen (1997). Stassen (1997), perhaps the most impressive result in comparative grammar to date, compares 410 languages on the point that concerns Becker. His conclusions, though, seem to be the exact opposite of Becker's conjecture. Stassen divides predicative constructions in four semantic types, each one tends to be encoded by a specific category: event/state (encoded by V), class-membership (by N), locational predicates (by Adv), and quality (by Adj). Stassen argues that languages have at their disposal three strategies for encoding the four semantic types of predicate: the verbal, the nominal and the locational strategy. Each semantic type has, comparatively speaking, its own preferred grammatical form. (a) Inflection for the verbal strategy. (b) Zero encoding or encoding by a verbal or non-verbal copula, for the nominal strategy. (c) Encoding by a copula or 'posture' (locative) verb for the locational strategy. (d) Quality predicates have no prototypical encoding strategy. What is important in the present discussion is that, from a cross-linguistic point of view, Stassen formulates two striking universals for the grammatical encoding of predicates. Firstly, zero encoding of predication may turn up in various predicates (adjectival, nominal or locational), but if a language allows this option it always allows it at least for nominal predicates, and especially for identity statements. (Stassen 1997, 64;120): Secondly, locational encoding seems to have an overwhelming tendency to require the presence of a supporting 'posture' verb' (1997, 56).

The semantic distinction made by Becker between properties intended to be permanent and properties intended to be temporary returns in Stassen (1994, 1997). He uses the qualification 'more/less time-stable' for the Spanish/Portuguese *ser*/*estar* distinction and for the nominative/instrumental Cases in Russian. It is true, as Becker also observed, that some languages use zero marking for the less time-stable option and the copula for the more time-stable alternative, cf. Hebrew. However, exactly the reverse situation can be attested for several other languages, like Ancient Indo-European. The zero option in those languages encodes the more time-stable alternative (Stassen 1997, 211f).

In general, zero encoding of predicates seems to be a rather restricted option from a typological point of view. Nevertheless, it turns out that, 'if ' zero marking is applied at all, it is mostly the inherent <–Asp> predicate that needs no time qualification in principle and applies zero marking. This is the exact opposite of Becker's UG conjecture. The picture Stassen gives is much more complicated and subtle than I sketched here, but I will not pursue the matter further at the present moment.

5.   Conclusion

The hypothesis of a universal innate multiple I-structure has had a considerable impact on recent work in language acquisition (Hoekstra and Hyams 1998, and others) and it is likely to be around for some time. Becker's thesis links up with this general idea of 'abstract syntax'. I have suggested more caution for acquisition research. Acquisition research may rather show that the learner never considers UG possibilities that are not supported by massive PF input evidence. The fact that evidence for a universal innate structure is easily overestimated, does not distract from the fact that Becker's thesis has been a courageous attempt for a detailed study of copula *be* constructions in child language. Had she not offered the detailed lists with numbers in clear oversights, I would not have started to construct graphs and to doubt the actual significance of that very UG factor.

References

Carlson, G. (1977). References to kinds in English. PhD. dissertation, University of Mass., Amsherst.

Diesing, M. (1992). *Indefinites*. Cambridge, Mass.: MIT Press.

Heycock, C. (1995). The internal structure of small clauses: new evidence from inversion. *Proceedings of NELS* **25**, J. Beckman (ed.). Amherst, Mass: GLSA.

Hoekstra, T. and N. Hyams (1998). Aspects of root infinitives. *Lingua* **106**, 81-112.

Jordens, P. (1990). The acquisition of verb placement in Dutch and German. *Linguistics* **28**, 1407-1448.

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

Kratzer, A. (1995). Stage-level and individual-level predicates. In *The Generic Book*, G. Carlson and F. Pelletier (eds.). Chicago: University of Chicago Press, 125-175.

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

Schlichting, J.E.P.T. (1996) *Discovering syntax*. Nijmegen: Nijmegen University Press.

Stassen, L. (1994). Typology versus mythology: the case of the zero-copula. *Nordic Journal of Linguistics* **17**, 105-126.

Stassen, L. (1997). *Intransitive predication*. Oxford Studies in Typology and Linguistic Theory. New York: Oxford University Press.