Locality and Evidence Frames as learnability tools for language acquisition

Jacqueline van Kampen
UiL OTS, Utrecht University
Trans 10
3512 JK Utrecht, Netherlands
Jacqueline.vanKampen@let.uu.nl
http://www.let.uu.nl/~Jacqueline.vanKampen

Arnold Evers
UiL OTS, Utrecht University
Trans 10
3512 JK Utrecht, Netherlands
Arnold.Evers@let.uu.nl

Abstract

Chomsky (2001b) points out that the properties of grammar must be explained from:

(i) primary linguistic data in child language.
(ii) an assortment of human perception and memory talents that co-operate in a language acquisition procedure
(iii) constraints that derive from quite general neural mechanisms

The present paper is concerned with primary linguistic data (i) and their interaction with the acquisition procedure (ii). The initial child data lead to a highly simplified grammar that nevertheless directs the further options of the acquisition procedure. The interaction of the factors (i) and (ii) can be studied by a closer attention to the order of acquisition steps. The relevance of (iii) is less clear. Locality and inclusiveness as general properties of grammatical structures rather seem to provide learnability without revealing much of the neural system as such.

The first half of the paper argues that the acquisition procedure starts with a radical reduction of the input data. This results in a sequence of evidence frames. The Single Value Constraint (Clark 1992) and the determined order of acquisition steps are explained when each acquisition step is related to its own evidence frame.

The second half of the paper demonstrates how the I(nflection)-marking of predicates and the D(eterminer)-marking of arguments are successive steps. Each invests the marked phrases with properties known as EPP and UTAH. These principles derive from the input. Once acquired, they guide the acquisition procedure to – for example - dummy subjects and agreement, rather than the other way around.

1 Input Reduction

The universal locality restrictions in human grammar determine the way language acquisition operates. This is not meant as a particularly profound insight. There is very little else for an initial acquisition procedure to adhere to. Structures for stress rules, abstract categories, or empty categories, let alone underlying representations and transformations may be innate or not, they cannot be applied until there is some grammar. Initially of course, there is no grammar yet, since the acquisition procedure is still on its way to find one out.

Longitudinal analyses of child language show how the child adds grammatical markers within reduced binary frames and how each syntactic acquisition step relies on such a preceding binary frame. The locality frames in adult grammar testify of a prehistory in which they were the origin of an acquisition step, i.e. the rise of a new functional feature. This perspective on syntactic locality is given in (1).

(1) a. all acquisition frames, c.q evidence frames, are based on a local adjacency in binary structures.
b. all marking by grammatical features is acquired with respect to such acquisition frames/evidence frames.

An evidence frame is defined as in (2)

(2) *Evidence Frame*
An evidence frame is a binary construction that can be fully interpreted by the current grammar but for a single functional category <F?>

The acquisition frames indicate how grammar is caused by input. The order of learning steps demonstrates that typological properties are the first ones to be acquired (as predicted by Jakobson, 1942). Abstract grammatical principles may result from a self-organizing learning procedure. Universal Grammar (UG) and its typological effects may be an outcome of the acquisition procedure rather than its source.

The procedure for first language acquisition is not confronted with all possible grammatical problems at once. This is not even possible. The somewhat more complicated problems cannot become visible unless distinctions are made between several domains and categories. Longitudinal graphs from child data show that grammatical principles *fade in*. They appear over time, take a few months to get established and moreover the order of acquisition steps follow a logic that can be explained. Acquisition order follows a grammatical hierarchy.

Before discussing whether some grammatical property can be acquired, one must agree upon which part of the grammar is already available to the learner. We propose that the acquisition procedure applies at first considerable reductions on the observation space (input data). A main point is the distinction between *input* and *intake* (White 1986). The mother's input is reduced to a radically different intake by the child. The reduction to the intake follows from a simple common sense principle, see (3).

(3) *Reduction of input to intake*
Leave out what you cannot fit in and try minimal solutions.

Suppose the acquisition procedure starts with the reduction operation in (4)

(4) *Input Reduction*
a. substitute <+F,?> for each grammatical marking that is still unknown.
b. throw out all input sentences with more than one <+F,?>.
c. attach Fi as a marker to the selectionally dominant element to left or right

The result is an intake that satisfies the *Single Value Constraint* (Clark 1992). The intake to the acquisition procedure under these restrictions is now such that one grammatical category is singled out, identified and subsequently acquired. The acquisition of the category $<+F_i>$ changes the data reduction procedure. The next grammatical category $<+F_{i+1}>$ is singled out, etc. Below, in section 4, we will present acquisition data that support the reductions in (4).

2 Input-control

It has never been disputed that language acquisition depends on a certain amount of input-control. Yet, it has so far not been a research priority in generative grammar to see how the appropriate input could be selected, what pre-existing grammar it could affect, and whether the order or the speed of the acquisition steps could be predicted. Generative learnability theories in the 1980th were theoretical and somewhat defensive. They qualified the mathematical deduction in Gold (1967) that context-free rewriting grammars could not be identified or learned without negative data. As Wexler & Culicover (1980) argue, context free generative grammars and some transformational grammars are learnable from positive data as long as the relevant relations are sufficiently local. The main point was to argue a learnability in principle for certain types of generative grammar. There was no reference to child language.

The ongoing simplification of grammatical principles, pushed by Categorial Grammar, HPSG and the Minimalist Program, may re-inspire interest in their learnability. Let's compare four attempts into that direction. Fodor (1998, 2001), Yang (2002), Culicover & Nowak (2003) and our own work (Van Kampen, 1997; Evers & Van Kampen, 1992, 2001). Fodor (1998, 2001) and Yang (2002) assume that the child is confronted with the full variety of constructions in his language. The child meets this challenge with brilliant creativity. She comes up with all possible grammatical structures that the general theory of grammar would allow. The child's productivity in designing possible solutions is maybe comparable with his creativity in grasping visual or musical structures or maybe with the babbling phase that precedes the construction of phonological forms. Fodor as well as Yang's learner start with a variety of grammatical structures and work towards a minimal set of grammatical structures. Fodor's learner is sensitive to certain key-constructions (treelets) that betray the language type and Yang's learner is sensitive to rules that are too often involved in analyses that fail. Yang proposes an accounting system of "penalties" for failing rules.

Yang's bookkeeping of failures and Fodor's testing system could be characterized respectively as an effective *evaluation procedure* (Yang) and as an effective *decision procedure* (Fodor). They, as well as their little learners, start with all options offered by the theory. Subsequently, they propose computational operations that select a language-specific grammar for the input data. Both successfully simulate how the learner zeros in on the core grammar of the language.

This is not the only approach possible. We propose, like Daelemans et al. (2000), Culicover & Nowak (2003), that the young learner is unaware of the grammatical alternatives that are available in the world outside. Our learning procedure could be characterized as a *discovery procedure*.[1]

[1] See Chomsky (1957:50ff) for the distinction

Suppose the child has reached a point at which she is able to recognize separate words. Then the first input reduction will be that the child ignores all grammatical marking, like articles, auxiliaries and verbal inflections. Functional categories characterize syntactic combinations and they cannot be identified and learned until there is lexical content material that is understood as a combined expression. What remains is a limited set of lexical content words that are learned as names and characterizations in actual situations. *Ain't the bear nice? The bear is nice. I want the bear to be nice.* are all turned by the child into the single [*bear nice*]. A set of binary constructions is the result.

The acquisition order is due to input-control, but definitely not always due to input frequency. For example, functional categories are acquired later than content words, yet functional categories are more frequent. Their token frequency is 100 to 300 times higher than the token frequency of an arbitrary content word. Although highly frequent, functional categories can be learned only in constructions that contain content words. The acquisition order "content words before functional categories" is imposed by the nature of the system the child is confronted with. Not by an innate grammatical principle.

Eventually, the learning procedure identifies grammatical markings between binary combinations of content words one at a time. This is a possible characteristic of actual learning. It has been studied in formal learnability theory as the Single Value Constraint (Clark 1992). We hope to derive the Single Value Constraint from the successive reductions on the input. The fact that grammatical words and endings are left out in the child's early productions is less prominent in Fodor (1998, 2001) or Yang (2002). Their learners remain in direct contact with the full input. By contrast, our young learner must reduce its initial attention to constructions assigned to pairs of adjacent content words and so he enters a maximally reduced observation space, as predicted in (4).

3 Locality and blocking

Language acquisition must overcome its radical underspecifications. It proceeds by adding grammatical features within a local binary frame (Van Kampen 1997). A learning procedure that adds a grammatical feature to a category moves from a less restricted superset to a more restricted subset. The learning procedure starts with underspecifications, but the associative pressure of local contexts has a healing effect. The initial underspecifications are "blocked". Blocking effects are known from the very beginning of grammatical studies (Panini, DiScullio & Williams 1987). Irregular inflection forms are said to block the regular ones. The initial option that allowed *she gived the apple* and *she gave the apple* is reduced to the latter (Marcus et al. 1992). The acquisition procedure eliminates the option between a less specified and a more specified variant. In general, the more specified variant blocks the less specified one.

The acquisition procedure will add the lexical specifications as grammatical features if they are present in the input sufficiently early and sufficiently robust. Blocking is a procedure over time and the more specified variants compete for some time with the earlier and less specified variants. Blocking is more an effectiveness device. It never works instantaneously. It takes some time and some quantification before the learner reacts. This reminds of Yang's (2002) penalty system, but Yang's system is more informed and intelligent. It chooses between innate alternative grammatical solutions. Our system is more stupid. It is pressured by mere frequency to add grammatical specifications to an underspecified frame. The underspecified more general form is subsequently blocked (Van Kampen 2004b).

The blocking procedure is also effective outside morphology. For instance, the Dutch input offers the young learners about 25% VO (Verb-Object) patterns and about 75% OV (Object-Verb) patterns. The latter is largely due to a quantitative amount of auxiliaries and modals in the input. At a very early stage of development the 75% OV pattern wins the competition and the VO patterns get marginalized. The lexicon of content words establishes a strong preference for OV patterns. (Evers & Van Kampen 2001).

The natural locality restriction in acquisition allows the underspecifications followed by blocking. The domain restriction itself is an immediate safeguard against domain over-generalization. Grammatical specifications like reflexivization and wh-movement in Dutch child language will not appear before there are finite verbs and finite verbs will not appear regularly before theta frames are established (Van Kampen 1997). Subject-verb agreement, reflexivization and scope of wh-movement are learned with respect to the local IP. The later embedding of for example an infinitival IP within a matrix IP, does not alter the strictly local character of the IP internal markings that have been acquired before.

between discovery, decision and evaluation procedure.

4 Order of acquisition steps: I°/predicate-marking precedes D°/reference-marking

Blocking in language acquisition can be traced by longitudinal graphs. The acquisition graph reflects "parameter setting" on the more specified value. See the graphs in (5) for Dutch I- and D-marking (figure 1) and for French I- and D-marking (figure 2) (Van Kampen 2004c).

(5)

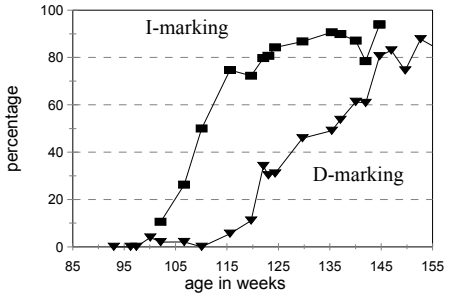


Figure 1: Dutch Sarah

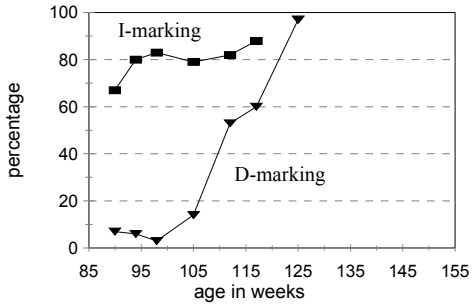


Figure 2: French Grégoire

Now the order of acquisition steps shows that both Dutch Sarah and French Grégoire apply systematic I-marking almost half a year earlier than systematic D-marking. Sarah has acquired I-marking at week 120 and D-marking at week 145. Grégoire has acquired I-marking at week 94 and D-marking at week 125. The same order of appearance was found for Rumanian (Avram & Coene (2004).

The amount of determiners (articles/ demonstratives) outweighs the amount of copulas/auxiliaries in the input data. Yet, children in various languages start to analyze predicate-argument structure by I-marking. The less frequent I-marking precedes the more frequent D-marking in acquisition. The acquisition procedure appears to follow the *Single Value Filter on Evidence Frames*. Initially, sentences with both a D-marked noun and an I-marked verb are thrown out of the observation space. D-marking, although more frequent, will not often occur without an I-marking {*the bear must eat*}, whereas I-marking may and will often occur without a D-marking {*bear must eat*}. If so, "I-marking precedes D-marking" is potentially a matter of universal acquisition order.

The systematic I-marking and D-marking themselves give entrance to a whole series of further acquisition steps, beginning with a grammatical decision procedure on the category membership V versus N (Van Kampen 2005). This option, chosen here for language acquisition, was implemented earlier in computational approaches to category assignment (Buszowski 1987).

4.1 Acquisition steps due to local evidence frame IP/DP

All the primary steps are a feature spelled out in a context of adjacent sisters.

Step 1 *I-marking and the category <+/– V>*
The systematic marking of "comment" parts is realized by a variety of devices {copula, modal, auxiliary, inflection} The distinction between the lexical heads in the predicate follows from their association with different I°-markings as in (6).

(6) $X^o \rightarrow$ <+V> / I^omodal —
$X^o \rightarrow$ <+V> / $I^o_{<inflection>}$

$X^o \rightarrow$ <–V> / I^ocopula

There are no mistakes in the selection of I° and <+/–V°> items. This shows that acquisition by repetitive context is an effective construction device. The short sentence forms used by children will enhance this effect (cf. Elbers 2002).

Step 2 *EPP (subject-requirement)*
Utterances consisting of no more than a comment are quite possible, but the I°-marked comments will rarely miss a topic. We intend to develop this into an argument that I°-marking is acquired as marking the context of a topic.

(7) ∅ → I^o / topic [— [comment]]

The acquisition of the I-marking is probably part and parcel of the acquisition of the EPP (subject-requirement). The standard EPP is the first step and the basic step in acquisition. The EPP is the acquisition step that must guarantee that each predicate is "anchored" due to the obligatory presence of a subject. It should be possible to demonstrate this quantitatively for any language.

Step 3 *D-marking and the category <+N>*
The naming topics that are not proper names are either a demonstrative or a D-marked noun. The D-marking of the topic by articles and attributive

pronouns can be captured in an acquisition graph (as in figure 3 below).

(8) ∅ → D^o / [— [X]$_{naming}$]

It stands to reason that the systematic association of certain lexical content words with D-marking, leads to the category N, see (9).

(9) X^o → <+N> / D^o —

The context of a sister D^o would not hold in the adult language, but within the simplified context of child language, there is sufficient evidence to figure out the category <+N>. By itself, the <+N> is not referential, but in a highly repetitive context D^o, it fits the same topic position as proper names and independent demonstratives. The acquisition rule may have the form in (10).

(10) ∅ → D^o / [— [X]]$_{referential}$

Step 4 *Free anaphors (clitics and pronouns)*
D-marking is not the marking of a somewhat hidden category <+N>. In the same vein, I-marking is not the marking of a somewhat hidden category <+V>. I-marking and D-marking are rather the grammatical expression for the predicative (characterizing) and the referential (naming) function of lexical content elements. The strange thing is that many positions where the X<+N> appears {*read booklet*, *drives car*, *sees house*} get the D-marking inserted. Personal pronouns (Dutch *hij/zij/haar/hem*; French *il/elle/lui*; English *he/she/it/him/her*) appear in the same frames at the same moment. The use of personal pronouns in the files of Dutch Sarah rises to input level simultaneously with the insertion of D^o elements next to the [X<+N>]. See the graphs in (11) figure 3. They underline Postal's (1966) view that pronouns and articles are both D^o.

(11)

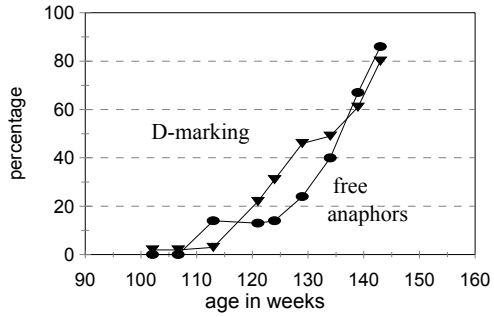


Figure 3: Dutch Sarah

This suggests strongly that the real acquisition step has been "mark naming elements by D^o". The naming parts are explicitly getting the referential function: Find out what is named by looking at the syntactic structure. Whereas naming and commenting are merely pragmatic (situation-bound) intentions, reference and predication are tied up with explicit grammatical marking. They cannot do without syntax.

The arguments establish the UTAH (Baker, 1985), i.e. the assignment of theta-roles to fixed subcategorizing positions. The UTAH offers an evidence frame for the subsequent acquisition of clitic arguments in French (Van Kampen 2004c). What we have in mind here is that D-marking is part of the acquisition of UTAH, like I-marking is part of the EPP (Van Kampen, 2004a). They stand for the acquisition of predicate structure (I-marking) and argument structure (D-marking).

4.2 Local evidence frames outweigh mere input frequency

A general property of 'decoding' emerges as well. The successive evidence frames narrow down to a far more precise context and the speed of acquisition increases by an order of magnitude.

The subject/topic of the previous steps still lacks φ-features of person/number (i.e. it is 3rd person singular only: Benveniste, 1966). D-marking establishes the EPP as relation between a head and its specifier. That opens the way to figure out in a subsequent step the φ-feature content in D^o, {± person, ± number} on the subject, i.e. oppositions between 1st/2nd/3rd person and singular/plural subjects.

(12) des motos (plural) fait (singular) du bruit
(motors make noise; Grégoire 2;01.25)

The finite verb still doesn't show the correct agreement with the subject. Late acquisition of agreement has also been reported for by Ferdinand (1996, for French), Schütze (1997), Avram & Coene (2003, for Rumanian). One step later, the initial I-marked predicate constitutes the local evidence frame for Agreement features, the copying of the φ-features on the I^o. The finite verb starts showing the correct agreement.

(13) elles sont aux Etats-Unis, mes sandals
(they are in the USA; Grégoire 2;05.23)

Now, both Sarah and Grégoire acquire φ-oppositions on the subject before the finite verb starts showing correct agreement

. The succession of the acquisition steps also shows the same relative speed. The later steps are a

matter of weeks whereas the earlier steps were a matter of months, see (14) for Dutch Sarah.

(14) *EPP as evidence frame*

step I^o	step D^o	step D^o (φ)	step I^o (φ)
20 wks	25 wks	5 wks	5 wks

The more effective acquisition relates plausibly to the more precise frame that can be used to select the input. The selection of some binary combination of content signs is far more undetermined than the distributional relation between explicit grammatical markings such as φ-features and Agreement. The later set of acquisitions is supported by a lexicon with categorial marking <+I> or <+D>.

We propose that after step 1 and step 2, the EPP operates as an evidence frame.

The input has not been lacking in φ-features on I and D, rather the φ-features could not become part of the intake before I and D had been established. It is only after the acquisition of I-marking and D-marking that the EPP begins to function as an evidence frame, a preceding structure that is needed to spot the relevant points. The effect of *evidence frames outweigh mere input frequency*. The specifier-head relation is acquired before the agreement marking. The latter appears as a final touch rather than a structural underpinning. One may also have notice that the addition of new material may take place at the outside of a frame as well as at at the inside.

5. Movement structures

None of all this implies a movement rule. The learnability of movement rules is a problem when movement rules are seen as rules that reorder an underlying array of heads and phrases in order to arrive at the perceived surface structure. At least two problems in (15) have then to be dealt with.

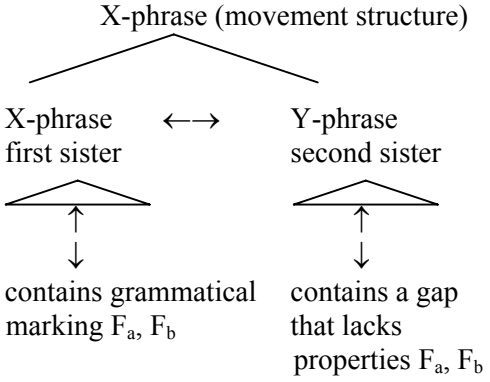
(15) a. *The gap* problem
What brings the learner to consider phrases as reordered, i.e. how can a phrase position be perceived as an antecedent or a gap?

b. *The distance problem*
How are syntactic 'islands' learned, i.e. how much distance can come between gap and antecedent?

Underlying structure is not audible, whereas surface structure is audible. It is not immediately clear how the perceived structure should lead to a rather different structure that is not perceived. Moreover, a free option "move category" must be reigned in by island constraints. These islands seem to ask for negative evidence rather than for positive examples.

Both problems are more manageable in unification-based approaches that trade in the movement rule for a lexical feature matching between two sister constituents (Neeleman & Van de Koot 2002). The first sister is grammatically marked, but out of context. For example, wh-phrases in the Spec,C position are case and preposition marked as if they held an argument position. In the same line, the finite verb in the C^o position carries the tense/agr markings as if it were in the I^o position. The second sister of the construction should contain a grammatically definable gap that would fit the grammatical properties of the first sister, see (16).

(16)



Obviously, the learnability of islands is on a promising track when the learner does already command a grammar that:

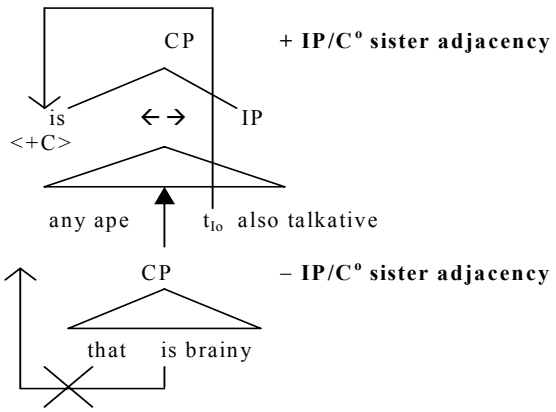
(17) a. spots the grammatical properties F_a, F_b orphaned in the first sister and lacking in the gap of the second sister, according to a current grammar.
b. projects the relevant grammatical properties F_a, F_b, etc., towards the phrase labels of the two sisters, according to existing conventions.

The learnability of the antecedent/gap status (15)a and the learnability of their legitimate distance (15)b is now shifted. Fortunately, the grammatical properties F_a, F_b, etc. that define the antecedent/gap relation have been acquired earlier in non-gapped structures. This is an empirical point and it fits (17)a. In addition, the grammar does already contain a procedure to match grammatical features. It was acquired when heads were subcategorized for grammatical properties of their complements. This fits point (17)b. The feature-

matching procedure works for subcategorized complements, but not for subjects and adjuncts. The latter, subjects and adjuncts, happen to be islands, whereas the subcategorizing complements are in principle non-islands. It is a remarkable effect that the wh-structures in child Dutch succeed to disrupt existing patterns that have grown over a year in a short period of less than 10 weeks, immediately after D-marking.

The learnability of island constraints has been has been a topic of debate (Crain & Nakayama 1987, *Linguistic Review* 2002). In (3), the copula from the main clause is fronted. Copula-movement out of a subject relative like (18) is not possible (Complex NP Constraint).

(18) Is_i any ape that is brainy t_i talkative?



The question 'how children would learn this structure dependant rule, ignoring the linear dependency' is related in Crain & Nakayama (1987), Yang (2002) to innate principles (UG). Pullum & Schulz (2002) related it to percentages in the input. In the present view, it is neither a consequence from UG nor a consequence from frequency in the input, but a consequence from feature projection (inclusiveness and sisterhood). The properties of the embedded copula will not reach the matrix projection line. *The ape* with subject status cannot do that. For that reason, preposing of the downstairs copula cannot be acquired. The grammar that has been acquired cannot figure that out, because there is no appropriate feature projection to lift the structure into comprehensibility.

To sum up, the evidence frames for acquisition steps are local relations between two sisters, one marked by a functional category Fi. A position within an island cannot appear as sister of a position outside of the island. hence, the two cannot be involved in the same rule. Such a rule is unlearnable. That's all. The positive requirement (sisterhood) is a quite natural one and it appears in each acquisition step again. Evidence for the <+IP/C^o> adjacency in (18) has been massive and unexceptional during acquisition. The most trivial relation is a natural rather than a genetic option.

6. Conclusion

We reached the following conclusions.

(19)

a. The order of acquisition steps can be recorded by longitudibnal graphs (fact)
b. The locality and inclusiveness properties remain present because they allow the learnability of the system (conjecture)
c. Universal and typological properties may have been selected in grammars for the high learnability of their evidence frames (conjecture)

References

Adriaans, P.W. & M.M. van Zaanen. (2004) 'Computational grammar induction for linguists', ms. University of Amsterdam and Tilburg.

Avram, L. & M. Coene (2003) 'Why is it difficult to reach Agreement?', in A.T. Pérez-Leroux & Y. Roberge (eds) *Romance Linguistics. Theory and Acquisition*, 247-262. Amsterdam: John Benjamins.

Avram, L. & M. Coene (2004) 'What early clitics can tell us about early subjects', in: J. van Kampen & S. Baauw (eds) *Proceedings of GALA 2003*, 93-102. Utrecht University, LOT Occasional Series.

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

Benveniste, E. (1966) *Problèmes de Linguistique Générale*. Paris: Gallimard.

Buszkowski, W. (1987) 'Discovery procedures for categorial grammars', in: E. Klein & J. van Benthem (eds.) *Categories, Polymorphism and Unification*. University of Amsterdam.

Chomsky, N. (1957) *Syntactic Structures*. Den hag: Mouton.

Chomsky, N. (2001a) 'Derivation by phase', in: M. Kenstowicz (ed.) *Ken Hale. A Life in Language*, 1–52. Cambridge MA: MIT Press.

Chomsky, N. (2001b) 'Beyond explanatory adequacy', *MIT Occasional Papers in Linguistics* 20 Cambridge MA: MIT Press.

Clark, R. (1992) 'The selection of syntactic knowledge' *Language Acquisition* 2(2), 83-149

Culicover, P. & A. Nowak (2003) *Dynamical Grammar: Minimalism, Acquisition and Change*. Oxford: Oxford U.P.

DiScullio, A.-M. & E. Williams (1987) *On the Definition of Word*. Cambridge MA: MIT Press.

Elbers, L.B. (2000) 'An output-as-input hypothesis in language acquisition', in: P. Broeder & J. Murre (eds.) *Models of Language Acquisition* Oxford: Oxford U.P..

Emonds, J.E. (1985) *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.

Epstein, S.D., E.M. Groat, R. Kawashima & H. Kitahara (1998) *A Derivational Approach to Syntactic Relations*. Oxford U.P., Oxford.

Evers, A. & J. van Kampen (1992) *An acquisition procedure for the I-projection in Dutch and English*. ms. UiL OTS, Utrecht University.

Evers, A. & J. van Kampen (2001) *E-language, I-language and the order of parameter setting*. Uil OTS Working Papers 00105-S-S, Utrecht.

Ferdinand, A. (1996) *The Development of Functional Categories. The Acquisition of the Subject in French*. PhD dissertation University of Leiden.

Fodor, J.D. (1998) 'Unambiguous triggers', *Linguistic Inquiry* 29: 1-36.

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

Gillis, S., W. Daelemans, & G. Durieux (2000) 'Lazy learning: a comparison of natural and machine learning of stress', in: P. Broeder & J. Murre (eds.) *Models of Language Acquisition*. Oxford: Oxford U.P.

Gold, E.M. (1967) 'Language identification in the limit', *Information and Control* 10, 447-474.

Jakobson, R. (1942) *Kindersprache, Aphasie und Allgemeine Lautgesetze*. Uppsula University.

Jakubowicz, C., L. Nash, C. Rigaut & Ch. Gérard (1998) 'Determiners and clitic pronouns in French-speaking children with SLI', *Language Acquisition* 7, 113-160.

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

Kampen, J. van (2004a) 'The rise of predication by EPP', in: A. Brugos, L. Micciulla & C.E. Smith (eds.) *Proceedings of BUCLD 28*, 267-281. Somerville MA: Cascadeila Press.

Kampen, J. van (2004b) 'An acquisitional view on optionality', *Lingua* 114, 1133-1146

Kampen, J. van (2004c) 'Learnability order in the French pronominal system', in: R. Bok-Bennema, B. Hollebrandse, B. Kampers-Manhe & P. Sleeman (eds.) *Romance Languages and Linguistic Theory*. *Selected Papers from Going Romance 2002*, 163-183. Amsterdam: John Benjamins.

Kampen, J. van (2005) 'Language specific bootstraps for UG categories' *International Journal of Bilingualism*.

Marcus, G.F., S. Pinker, M. Ullman, M. Hollander, T.J. Rosen & F. Xu. (1992) 'Overregularization in language acquisition' *Monographs of the Society for Research in Child Development* 57.

Neeleman, A. & H. van de Koot (2002) 'The configurational matrix', *Linguistic Inquiry* 33-4, 529-574.

Postal, P. (1966) *On so-called pronouns in English*, in: F. Dinneen (ed.) *Monograph Series on Language and Linguistics* 19, 177-206. Georgetown: Georgetown U.P.

Pullum, G. & B. Schulz (2002) 'Empirical assessment of stimulus poverty arguments', *Linguistics Review 19* 1-2, 9-50.

Schütze, C. (1998) *INFL in Child and Adult Language: Agreement, Case and Licensing*, PhD dissertation MIT Cambridge MA.

Wexler, K. & P. Culicover (1980) *Formal Principles of Language Acquisition*. Cambridge MA: MIT Press.

Yang, C.D. (2002) *Knowledge and Learning in Natural Language*. Oxford: Oxford U.P.