Single value steps in first language acquisition
Jacqueline van Kampen and Arnold Evers

Abstract
Preferably, the properties of grammar can be derived from the following factors:
(i) The primary linguistic data as they are offered to the child.
(ii) A language acquisition procedure.
Hopefully, the language acquisition procedure is compatible with plausible assumptions about the neural abilities of human beings, but that is of no immediate concern here. The interaction of the primary data and the acquisition procedure can be studied by a closer look at the order of the child’s acquisition steps. What does the child acquire first and why? What does she acquire later and why? This is empirically a promising and by no means trivial approach. At the same time, we will argue against an assumption that is quite common in computational studies and also in mere grammatical studies of child language. People from Gold (1967) to Yang (2002) assume that the acquisition procedure has simultaneous access to all data at once. Our point will rather be that the acquisition procedure implies a natural selection of data (not based on UG assumption). Child language can be reconstructed as the result of successive steps in input reduction. The result is a series of successive intermediate grammars. Each grammar causes a new data-selection from the same type of input. Data-selection turns the poverty of a diffuse input stimulus into a focused stimulus by highly selective intake. Eventually, it may be shown that generative grammar is learnable without the postulation of grammatical a priories.

1 Input reduction
Early child language is not only stereotype and repetitive, it is above all simplified. There is quantitative difference with the adult input. There are less lexical items, shorter sentences and less grammatical markings. More importantly, there is a qualitative difference as well. Certain grammatical properties are not present at all in early child language. They will not appear until other properties have been acquired first. This order in acquisition steps, first studied in Brown (1973), must be explained. Child language can be represented as a series of grammars $G_i$, each one one step more specified than its predecessor until the series ends with the target grammar $G_n$ (Chomsky 1975: 119f).

\[ G_n \Rightarrow G_i \Rightarrow G_{i+1} \Rightarrow \ldots \Rightarrow G_n \]

Each step is an invention that makes sense due to the preceding grammar and due to a new selection of data from the input, see (2).

\[ G_i + D_{i+1} \Rightarrow G_{i+1} \]
The initial steps in the linear order offer the learner frames that determine the subsequent acquisition steps. The first steps are potentially relevant to all further steps and as such they are likely to have typological significance and to qualify as ‘macro-parameters’ (Baker 1996). For instance, the first evidence frames fit Greenberg’s (1966) local frames for language typology: VSO (e.g. Celtic), SVO (e.g. French), SOV (e.g. Dutch). The later steps take place in a grammatical environment that is far more determined by previous acquisition steps. One may be more inclined to think about ‘learning’ in case of the latest steps than in case of the first steps. For example, when the verbal paradigm has been fully acquired, one may think about the irregular verbs as innocent substitutions that finally block the use of regular ones by sheer repetition. The regular verbal paradigm, also a fairly language-specific affair, looks more learnable if the category Verb, the auxiliary-verb restrictions and the φ-features on the subject are already acquired. In the same way, it can be argued that the language-specific markings for predicate-argument structure in the reduced data induce grammatical principles like EPP (subject obligation for inflection-marked predicates) and UTAH (standardized theta roles for fixed arguments), see Van Kampen (2006). Along this line, an innate UG need no longer be the source of parallels between grammars. The instructive force of input data is enhanced by the input reduction. Elementary distinctions are acquired by the force of an elementary reduced data set and child language itself testifies to the data reduction that is at issue. UG rather happens to be an outcome of an uninformed acquisition procedure.

It is uncontroversial that the child begins its acquisition process with a radical reduction of the mother’s (adult) input. The initial questions are formulated in (3).

(3) Initial research questions
   a. How does the child succeed to simplify the structures that she hears in a sensible way?
   b. Why are certain reductions repaired early and others later?

The first question (3)a can be answered by the provisionally simplified common sense principle in (4). We will give it a more subtle form in (8) below.

(4) Principle for input reduction
   Leave out what you cannot fit in.

The second question (3)b can be rephrased as in (5).

(5) How does a grammar Gi select the data Di+1 in order to reach the next and less reduced form Gi+1?

Note that the relevant data Di+1 are not covered by Gi itself. The data Di+1 are simply hidden in the mass of other unanalyzed input data. For that reason, we might say that
the elements in $D_{i+1}$ still are a poor and diffuse stimulus to reach $G_{i+1}$. How does the acquisition procedure get them out? Let us consider a simple case of reduction first.

Grammatical words or markings may be hundred or more times as frequent as any of the denotational (content) elements, yet they cannot be acquired without a grammatically interpreted context. This is because a grammatical element $F_i$ indicates a grammatical relation between two phrases $[XP [F YP]_i]$. It is a word that carries no meaning beyond the syntactic relation. The word *and* does not mean ‘pair’, the word *but* does not mean ‘contrast’, the word *is* does not mean ‘property’. They express a certain relation between denotational terms $X$ and $Y$.

When all grammatical elements are left out, pragmatic understanding of the utterances is the remaining option and binary constructions of denotational elements are best understood in a pragmatic situation-bound way. Early child language does indeed show something like that. The residue in child language is basically a set of binary constructions that lack grammatical elements. What remains is a set of denotational and name-like words like $\{\text{dance/daddy/shoe/milk/nice/warm/off/on/gone}\}$ and a set of pragmatic operator-like words $\{\text{no/more/this/that/wanna}\}$. Denotational elements are acquired before grammatical elements, although each of the grammatical markings is several hundreds times more frequent. The acquisition order is imposed by the nature of the system the child is confronted with. Denotational words can be acquired without the support of a grammatical context.

(6) *Typical reductions in early child language*

<table>
<thead>
<tr>
<th>utterance</th>
<th>naming topic</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (the)  bunny  (is) dance(ing)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (I keep the) milk  (is) warm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. (my) shoe  (goes) off</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. daddy  (has) gone</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>operator</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>c. (I) wanna play</td>
<td></td>
</tr>
<tr>
<td>d. (you like) more milk</td>
<td></td>
</tr>
<tr>
<td>e. (I put) this on</td>
<td></td>
</tr>
<tr>
<td>f. (that is) no(t) nice</td>
<td></td>
</tr>
</tbody>
</table>

The initial denotational elements do not carry by themselves a referential or a predicative intention, nor are they marked by some $I^*$ (inflection) for predication or some $D^*$ (determiner) for reference. Such intentions are for the sympathetic listener to find out. This is the period that Lyons (1979) had in mind when he suggested that child language might have proto-predication as a forerunner of predication and also
that child language might have proto-reference as a forerunner of reference. We propose the interpretation of Lyons (1979) in (7).

(7) a. proto-reference ~ naming function
    b. proto-predication ~ comment/characterizing function
    c. proto-ilocution ~ (pragmatic) operator function

The functions of naming and commenting are pragmatic and not yet grammatically marked. The same denotational meaning X can in principle be used for naming and for commenting alike, as one can see in the examples in (6)b,d. Because there is yet no grammatically marked relation, there need not be a distinction N/V either. One may assume that the child as well as any animal can make a distinction between things, events, properties, states and what not. It is hard to think of any consciousness at all when such distinctions are not available. The grammatical input reduction only means that such cognitive distinctions have not yet been related to fixed phonological forms. The point is that proto-grammar does not translate these cognitive qualities into grammatical forms, whereas the later grammars do. The proto-grammar seems a good candidate for G in (1). It lacks any grammatically expressed relation between the denotational elements.

Since there is no grammar yet, the grammatical elements can be noticed as highly frequent, but they cannot yet be interpreted. Let such unknown, but highly frequent, elements be represented by <F?> (unidentified functional feature). It intends to cover any grammatical marking, whether by a separate grammatical word (is/do/the/a) or by a morphological affix (-ing/-ed/-s). The acquisition procedure may now start with the reduction operation in (8).

(8) Reduction procedure
    a. reduction: substitute <F?> for each grammatical marking still unknown
    b. intake: throw out all input sentences with more than one <F?>

A neuro-linguistic interpretation of the shift towards <F?> → <Fi> is present in Locke (1997 and references therein). Grammatically marked language coincides with activities in the left-brain hemisphere. The child’s early and grammatically still unmarked sentences, by contrast, correspond mainly with activity in the right-brain hemisphere. This tallies well with the affective and social function of one- and two-word utterances. Affective and social values are better represented in the right hemisphere. It now happens that the acquisition of formal grammar, the early grammaticalization between the second and the third birthday, correlates with a shift in the major brain activity from the right to the left hemisphere. Locke (1997) explains this shift in the following way. Both brain sides process the incoming signals. The affective and social values of the signal provoke activity in the right hemisphere. Initial right hemisphere dominance for such values is thereby enhanced. The grammatical markings, e.g. for l’ (finite verbs) and D’ (articles), are noticed only as mere phonological particularities, not reacted upon and left out in the pragmatic
exercise of the right hemisphere. For that reason, the grammatical markings (although uninterpreted) are better noticed by the less involved left hemisphere.

The cognitive coherence between the meaningful denotational words begins with a limited number of stereotypes (Tomasello 2003), but when new denotational words stream in and the number of possible combinations increases accordingly, the holistic and cognitive understanding of the right hemisphere gets more difficult and slows down. As soon as the highly repetitive left-hemisphere phonological structures (order, stress, grammatical markings) begin to be recognized as schemes for coherence, there is a shift from right to left. The left-hemisphere phonetic image offers the abstract grammatical orientation for predication and argument structure. The acquisition of I° and D° offers a new entrance to the lexicon. This new entrance is far more effective. The grammatical orientation propels the child subsequently to a further and ultimately more than tenfold higher extension of her lexicon. It is a shift from proto-grammar towards real grammar.

The topic-comment relation is underlined by a functional element. When the functional element is reinvented as marking the topic-comment function, one may say that the comment gets the (grammatical) predicate function and that it implies the subject/topic. What we have in mind is that all functional categories are acquired in that manner. They should all originate from a binary frame X and Y that is related by an initially not yet understood <F°>. Let us call such frames ‘evidence frames’.

(9) Evidence frame
   a. pragmatically: an intuitively understood utterance
   b. syntactically: a binary phrase structure [ XP [ F° YP ]FP ]FP
   c. semantically: fully interpretable but for a single <F°>

The result of the input reduction will be that the intake to the acquisition procedure is a set of repetitive short sentences all containing a functional marking in the same position. The evidence frame may be represented as [ XP <F°> YP], where XP and YP consist of fully recognizable items or phrases. When it is possible to figure out what the syntactic form and meaning of the evidence frame is, we have a minimal acquisition step. The category <F°> is learned. The question mark disappears and the result is an identified grammatical category <F>, as in (10)a, exemplified in (10)b.

(10) Learning step
   a. identify the pragmatically understood <F°> as <F> and attach <F> as a grammatical marker to the selectionally dominant element to left or right.
   b. <FP°>  
      [Bear]  <FP°>  
      <F°>  YP  
      is  [nice]
The present example \([X \ [\text{Fi Y}]_\text{FP}]_{\text{FP}}\) needs additional motivation for why the functional category adjoins to \(Y\) rather than to \(X\). One of the reasons is that \([\text{Fi Y}]_{\text{FP}}\) appears in the proto-grammar as a pragmatic operator + denotational word, see (6)e-f. Other arguments may follow from selection relations.

Besides the [specifier \(X\)-head \(\text{Fi}\)] relation, the acquisition procedure will need evidence frames for complements, attributes, movement and scope-bearing elements. Suppose now that such evidence frames can be found for the respective functional categories/features. Then we may arrive in the following world. The evidence frames select the relevant data from the input, and they focus the acquisition procedure on the grammatical category \(<\text{F\?}>/\text{Fi}\) that relates to the binary evidence frame. The weak and diffuse stimulus that follows from the input sentences is now temporarily changed into a strong and exclusive stimulus that follows from the evidence frame.

The evidence frames are effective because they are minimal, but at the same time fully interpretable utterances. The relation between the parts \(XP\) and \(YP\) may have been recognized pragmatically as \(YP\) (nice) ‘characterizes’ \(XP\) (Bear). The label FP in (10)b implies that \(XP\) and \(YP\) are united by the relation \(<\text{F\?}>\). The binary structure, where \(<\text{F\?}>\) combines to the right with \([\text{nice}]\), is determined by the existence of the phrase type \([\text{is nice}]\) in proto-grammar and because it is further related to stress and selection patterns (Van Kampen 2005). The more precise construction of viable evidence frames will be spelled out in the book. The central point is that the in-between grammars \(G_i\) in (1) have a data-selecting effect that supports the learnability of \(G_{i+1}\). The basic claim that the intermediate (reduced) grammars select the data for the next acquisition step has been made before (Berwick and Weinberg’s ‘data-focus’ 1984: 284). Here it is used as an argument against the poverty of the stimulus. The reduction turns a diffuse stimulus into an effective one.

As one may realize, the reduction procedure applies recursively. When \(<\text{F\?}>\) has been identified as \(\text{Fi}\), a new grammatical feature has been acquired. By consequence, the input reduction in (2) will now reapply to the same kind of input, but deliver a new kind of intake. \(\text{Fi}\) now passes the reduction filter. The next grammatical category \(F_{i+1}\) is singled out, etc. The successive input reductions that follow in this way should remind us of Clark’s (1992: 90) ‘Single Value Constraint’ (cf. also Berwick 1985: 108). Notice though that the present acquisition procedure is not error-driven. It rather adds new specifications in a stepwise fashion. One may wonder why natural languages are so child-friendly to produce evidence frames that follow from systematic input reduction. A glib answer would be that grammatical systems of a less manageable type would fail to be learned and so fail to survive. A more constructive perspective is possible as well, and this is the perspective we ultimately want to explore. Let there be a type of formal systems that can be decoded by systematic input reduction. The decoding should ultimately arrive at a hierarchy of functional categories and their associated phrasal types. Is it possible that human core grammars are of that formal type? Main properties of human grammars, such as \(\text{locality}\) and \(\text{inclusiveness}\) as viewed in present day Minimalism, are bound to be stable when they guarantee a natural learnablity of the system.


2 Order of acquisition steps

Language acquisition overcomes the radical underspecifications that result from the initial data reduction by adding grammatical features within a binary frame. The order of the successive acquisition steps can be shown by longitudinal graphs, as in (11).

(11) Dutch Sarah: acquisition of I*-marking and D*-marking

The graphs in (6) represent the acquisition of finite verbs (I*-marking) and determiners (D*-marking) by the Dutch child Sarah (Van Kampen 2004a). The graph for I*-marking shows the growing percentage of grammatical predicate marking, {copula/auxiliary/modal/finite morphology}. The graph for D*-marking shows the growing percentage of argument marking {article/demonstrative/possessor}.

Now the order of acquisition steps shows that Dutch Sarah applies systematic I*-marking almost half a year earlier than systematic D*-marking. The same order of appearance was found for English, French, and Romanian. The amount of determiners outweighs the amount of finite verbs 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 order I*\rightarrow D* must be explained. Evers Van Kampen (2006) show how the acquisition procedure follows the Single Value Constraint on evidence frames as proposed in (9), when initially sentences with both a D*-marked noun and an I*-marked verb are thrown out of the observation space. The feasibility of a mechanical reduction procedure was partly demonstrated by a computer simulation (Obdeijn 2004) that derived an order of intake frames in child language from a child-directed input.

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. This option, chosen in (Van Kampen 2005) for language acquisition, was implemented earlier in computational approaches to category assignment (Buszowski 1987). 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 of the I*-marked predicate (finite verb) initially lacks \( \varphi \)-features of person/number. In a subsequent step, the \( \varphi \)-feature content in D*, \( \{\pm \text{person}, \pm \text{number}\} \) on the subject, is
figured out. However, the finite verb still doesn’t show the correct agreement with the subject, see (12).

\[ 12 \quad \text{de clowntjes}^{\text{+plur}} \quad \text{heb}^{\text{+sing}} \quad \text{oogjes} \quad \text{Sarah week 130} \]

(the clowns has eyes)

One step later, the initial $\Gamma$-marked predicate constitutes the local evidence frame for Agreement features, the copying of the $\varphi$-features on $\Gamma^\circ$. The finite verb now starts showing the correct agreement. Late acquisition of agreement has been reported for various languages, e.g. for Rumanian, Portuguese and Catalan.

The dense succession of the acquisition steps shows that the later steps are a matter of weeks whereas the earlier steps were a matter of months (Kampen 2006).

\[ 13 \quad \text{step } \Gamma^\circ \quad \text{step } D^\circ \quad \text{step } D^\circ(\varphi) \quad \text{step } I^\circ(\varphi) \]

\[ 20 \text{ wks} \quad 25 \text{ wks} \quad 5 \text{ wks (?)} \quad 5 \text{ wks} \]

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 $\varphi$-features and agreement. The later set of acquisitions is supported by a lexicon with categorial marking $\langle +I \rangle$ or $\langle +D \rangle$. After step 1 and 2, the EPP (subject-$\langle +\text{fin} \rangle$-verb configuration) operates as an evidence frame.

3 A discovery procedure

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 and 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 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. I mention four attempts into that direction. Fodor (1998), Yang (2002), Culicover and Nowak (2003) and our work (Evers and Van Kampen 2001).

Fodor (1998) 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 by comparing alternative solutions. 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. The options that they compare are assumed to be a priori present from the human brain. 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). Their learners start with all options offered by the theory. Both successfully simulate how the learner zeros in on the core grammar of the input language.

By contrast, we propose, like Culicover and Nowak (2003), that the young learner is unaware of any grammatical alternative that is available in the world outside. Our learning procedure could be characterized as a discovery procedure. 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 formulated in (8).

A learning procedure as in (10) that adds a grammatical feature to a category moves from a less restricted superset to a more restricted subset (Kiparsky 2002). 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, Kiparsky 2005). This is a contentious issue in theories of language acquisition. Some try to reconstruct child language as subset language that is extended to the correct generalizations by positive data (conform the Subset Principle). Others believe that child language starts with overgeneralizations and narrows down by developing subcategories (Jakobson 1942; the present perspective: Van Kampen 1997, 2004b).

Blocking in language acquisition can be traced by longitudinal graphs as we have seen in (11). Blocking never works instantaneously. It takes some time and some quantification before the learner reacts. Blocking is more an effectiveness device. 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 input frequencies to add grammatical specifications to an underspecified frame.

4 The learnability of movement structures

The evidence frames are also effective for the learnability of scopal phenomena, like wh-marking and negation. 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, the learnability of movement rules offer at least the two problems in (14).

(14)  a. The gap problem  
         How can a phrase position be perceived as an antecedent or a gap?

       b. The distance problem  
         How are islands learned as phrases that do not allow a gap and an outside antecedent?
Both problems are more manageable in unification-based approaches that trade in the movement rule for a lexical feature matching between two sister-constituents. Neelkeman and Van de Koot (2002) derive such an approach from Minimality principles. The first sister is grammatically marked \(<F_a, F_b>\), but beyond the usual licensing context. For example, wh-phrases in the Spec,C position are case-marked and preposition-marked as if they held an argument position. In the same line, the finite verb in the C° position carries the tense/agr markings as if it were in the I° position. The second sister of the construction should contain a grammatically definable gap \(<+C-\text{gap}, F_a, F_b>\) that fails to carry the marking, \(<F_a, F_b>\) see (15).

\[
(15)\quad \text{\textit{(<+C> movement structure}}
\begin{align*}
&\text{X-phrase first sister} && \leftrightarrow && \text{Y-phrase second sister} \\
&\downarrow \text{with a marking for }<F_a, F_b> && \downarrow \text{with a gap }<+C-\text{gap }F_a, F_b>
\end{align*}
\]

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

\[
(16)\quad \begin{align*}
&\text{a. spots the markings }F_a, F_b\text{ orphaned in the first sister and spots their lack in the gap }<+C-\text{gap }F_a, F_b\text{ to the top label }<&+C>. \\
&\text{b. projects the markings }<F_a, F_b\text{ according to existing conventions.}
\end{align*}
\]

Fortunately, the grammatical markings \(F_a, F_b\), etc. that define the antecedent~gap relation have been acquired earlier in non-gapped structures. This is an empirical point and it fits (16)a. Does the grammar contain a procedure to project grammatical features? Sure, it was acquired when heads were subcategorized for grammatical properties of their complements. This fits point (16)b. The subcategorizing feature matching 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.

A grammatical complement property is reflected in a property of the head it is governed by, and [head complement] phrases that are complements themselves. This is equivalent with the more colorful description where a \(<+C/+\text{wh}>-\text{marked constituent is first present in its argument position, moves up to the }<+C>\text{ position, but requires a head~government path between }<+C>\text{ and the gap as proposed in Kayne (1981). In both arrangements (Kayne 1981, Neelkeman and Van de Koot 2002) the positions within a non-complement cannot be considered. This turns the notion island into a non-entity. It is not present in the grammar, hence it need not be learned.}

The learnability of island constraints has been a topic of debate (Chomsky 1975, 1980: 319, Crain and Nakayama 1987, Linguistic Review 2002). In (17)a, the copula from the main clause is fronted. Copula movement out of a subject relative like (17)b is not possible (subject island).
a. Is any ape that is brainy talkative?
b. *Is any ape that is brainy talkative?

It is a long debated issue how children learn that the movement is structure dependent, rather than linear dependent. Crain and Nakayama (1987), as well as Legate and Yang, Fodor and Crowther (among others Linguistic Review 2002) relate it to innate principles, whereas Pullum and Scholz (Linguistic Review 2002) prefer to see a hint in input percentages as sufficient. In the present view, it follows neither from innate principles, nor from input. It rather is a consequence from a feature projection system that has been acquired earlier (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.

The discovery procedure proposed here must assume that the natural input allows a reduction to local frames and a terminal string that remains informative enough in spite of the reduction. Locality and local inclusiveness of grammatical information are present to guarantee a certain type of learnability, without grammatical a priories.

References


