Formal Grammars of Early Language

Shuly Wintner[1], Alon Lavie[2], and Brian MacWhinney[3]

[1] Department of Computer Science, University of Haifa
shuly@cs.haifa.ac.il
[2] Language Technologies Institute, Carnegie Mellon University
alavie@cs.cmu.edu
[3] Department of Psychology, Carnegie Mellon university
macw@cmu.edu

Abstract. We propose to model the development of language by a series of formal grammars, accounting for the linguistic capacity of children at the very early stages of mastering language. This approach provides a testbed for evaluating theories of language acquisition, in particular with respect to the extent to which innate, language-specific mechanisms must be assumed. Specifically, we focus on a single child learning English and use the CHILDES corpus for actual performance data. We describe a set of grammars which account for this child's utterances between the ages 1;8.02 and 2;0.30. The coverage of the grammars is carefully evaluated by extracting grammatical relations from induced structures and comparing them with manually annotated labels.

1 Introduction

1.1 Language Acquisition

Two competing theories of language acquisition dominate the linguistic and psycho-linguistic communities [59, pp. 257-258]. One, the *nativist* approach, originating in [14,15,16] and popularized by [47], claims that the linguistic capacity is innate, expressed as dedicated "language organs" in our brains; therefore, certain linguistic universals are given to language learners for free, requiring only the tuning of a set parameters in order for language to be fully acquired. The other, *emergentist* explanation [2,56,37,38,40,41,59], claims that language emerges as a result of various competing constraints which are all consistent with general cognitive abilities, and hence no dedicated provisions for universal grammar are required. Consequently,

> "[linguistic universals] do not consist of specific linguistic categories or constructions; they consist of general communicative functions such as reference and predication, or cognitive abilities such as the tendency to conceptualize objects and events categorically, or information-processing skills such as those involved with rapid vocal sequences." [58, p. 101].

Furthermore, language is first acquired in an *item-based* pattern:

"[young children] do not operate on the basis of any linguistic abstractions, innate or otherwise. Fairly quickly, however, they find some patterns in the way concrete nouns are used and form something like a category of a noun, but schematization across larger constructions goes more slowly... The process of how children find patterns in the ambient language and then construct categories and schemas from them is not well understood at this point." [58, pp. 106-107].

Our ultimate goal in this work is to provide a formal environment in which these two hypotheses can be rigorously evaluated.

Children learn language gradually and inductively, rather than abruptly and deductively [37]. Still, it is possible to isolate "snapshots" of language development which are qualitatively distinct from previous snapshots, e.g., by moving from one word utterances to two words, exhibiting a new possibility for word order, introducing embedded clauses, etc. It should therefore be possible to construct grammars for each of these snapshots of language development in isolation, and then inspect the differences among those grammars. The underlying assumption of this work is that such snapshot grammars exhibit a gradual, smooth development which can be captured formally and accounted for mathematically.

As an example of "snapshots" of language acquisition, consider the first five *stages* defined by [11], namely:

1. Acquisition of *semantic roles* (e.g., agent, patient, instrument etc.) in simple sentences expressed by linear order and syntactic relations;
2. Acquisition of *semantic modulations* such as number, specificity, tense, etc., expressed morphologically or lexically;
3. *Modalities* of the simple sentences, e.g., interrogative, negative and imperative sentences;
4. *Embedding*, as in relative clauses or VP objects;
5. *Coordination* of full sentences or partial phrases with deletion.

[11] shows that these five stages correspond nicely to measures of grammar development, in particular mean length of utterance (MLU). However, there is nothing sacred in this number of stages and a finer granularity of representation is certainly possible. There is also evidence that in some languages (e.g., Hungarian) the order of the stages may be somewhat different [36]. A similar definition of "a phase-based model of language acquisition" is offered by [4], although the five stages she delineates are not identical to those of [11]. We will focus in this work only on the first 3–4 phases of [4], again dividing them up to a finer resolution.

In order to investigate the development of child language, corpora which document linguistic interactions involving children are needed. The CHILDES database [39], containing transcripts of spoken interactions between children at various stages of language development with their parents, provides vast amounts of useful data for linguistic, psychological, and sociological studies of child language development. Two developments make CHILDES an attractive experimental setup for investigating language acquisition. First, similar databases of

child-adult linguistic interactions are constantly being collected and developed for a variety of languages. Furthermore, many of the corpora are morphologically analyzed and annotated in a compatible manner, which makes it possible to compare language development across different languages. Second, the English CHILDES database has recently been annotated with grammatical relations [51]. This is useful for various practical applications (e.g., assessing the syntactic development of children, as in [52]), but is particularly attractive as it provides a way for automatically evaluating the coverage and, to some extent, the correctness of formal grammars which attempt to generate child language (see, e.g., [7]).

This paper presents some preliminary results. Following a discussion of related work below, we describe in section 3 a set of formal grammars accounting for the early utterances of one English speaking child, *Seth* [44,45], as reflected in the CHILDES database [39]. Faithful to the item-based model of language acquisition, the grammars that we develop are highly lexicalized (typed) unification grammars [13], inspired by Head-Driven Phrase Structure Grammar (HPSG, [48]). Such grammars provide means for expressing deep linguistic structures in an integrated way. We use LKB [19] to implement the grammars and apply them to the data in the corpus, producing deep linguistic structures which describe the syntax, but also some aspects of the morphology and the semantics, of child utterances. In particular, we convert the structural descriptions produced by the grammars to the functional annotation of [51]. This facilitates an automatic evaluation of the coverage of the grammars (section 4). On our development set, the error rate is lower than 10%.

1.2 Formal Grammars

Formal grammars are mathematical systems for describing the structure of languages, both formal and natural. In this work we employ *unification grammars* [65,66] and in particular unification grammars that are based on *typed* feature structures [13]. Such grammars underly linguistic theories such as Lexical Functional Grammar [28] and Head-Driven Phrase-Structure Grammars (HPSG) [48].

Typed unification grammars are based on a *signature* consisting of a partial order of *types*, along with an *appropriateness specification* which lists the *features* which are appropriate for each type and the types of their values. The signature facilitates the specification of an ontology of linguistic entities, from basic signs to rules and grammars. The building blocks of unification grammars are *feature structures*, which are nested data structures licensed by the signature. Grammar rules can then refer to and manipulate various pieces of the data encoded in feature structures, providing the linguist with powerful means for expressing linguistic theories in various strata: phonology, morphology, syntax, semantics and information structure. Unification grammars, and in particular HPSG grammars, have been used successfully for describing the structure of a variety of natural languages [3,26,29].

Typed unification grammars are attractive as a grammatical formalism, and in particular as an environment for investigating language acquisition, for several

reasons. First, grammars are multi-stratal, allowing for the expression of various types of linguistic information and interactions thereof in the same sort of object. Second, grammar rules are basically constraints which limit the domain of grammatical utterances rather than procedures for generating linguistic structures. Third, such grammars are highly lexical: most of the information is encoded in the lexicon and rules tend to be few and very general. This goes hand in hand with item-basd theories of language acquisition. Finally, the unification operation which is so essential to the formalism provides a natural implementation of unifying the effect of various competing constraints, although it currently does not clearly support the kind of violable constraints that are known in, e.g., optimality theory [49]. While unification grammars, as a constraint-based formalism, are adequate for characterizing *structure*, they may not be ideal for characterizing *processes*; as will be made clear in section 2, we focus in this work on the formal representation of linguistic knowledge, rather than on the processes which lead to its acquisition. Unification grammars, therefore, are a natural formalism to use for this task.

One task that unification grammars may fail in is the truthful modeling of *competition* among different, incompatible constraints. This seems to be one of the core mechanisms in language acquisition, and is probably best modeled with some kind of statistical or probabilistic framework. We conjecture that stochastic unification grammars [1,27] may be adequate for this, but we leave this direction for future research.

1.3   Formal Approaches to Language Acquisition

Many existing theories of language acquisition focus on exploring *how* language develops, applying computational learning theory to the problem of language acquisition [62,5]. They assume a class of formal languages and an algorithm for inducing a grammar from examples, and differ in the criterion of success (e.g., identification in the limit [22] vs. PAC learning [60]), the class of algorithms the learner is assumed to apply (accounting for memory limitations, smoothness etc.) and the class of formal languages that can be learned.

A different kind of formal approaches to language acquisition utilizes computer simulations; this line of research is sometimes called *computational approaches* to language acquisition [10]. For a discussion of such approaches and their interrelations with behavioral, experimental methods, see [9]. Of particular interest is the *What and why vs. how* dichotomy that [9] introduces: much of the research conducted under the computational framework seems to be interested in explaining *how* language (in any case, certain aspects of language) is acquired. In this work we are interested in how linguistic knowledge is organized and how this organization develops as the child matures; in the terminology of [9], this is a *what* question, which we nevertheless address formally.

The dominant nativist approach to language acquisition attempts to place it in the context of *principles and parameters* theories of language [14,17]. Assuming some form of Universal Grammar with a set of parameters that have to be finely tuned as a particular language is acquired, [21] propose a *triggering learning*

algorithm (TLA) which is aimed at setting the parameters correctly given positive examples. Some shortcomings of this algorithm were pointed out, e.g., by [43], who propose a model with fewer assumptions and better convergence performance.

In the main stream of generative approaches to syntax, a representative work on language acquisition is [31], which attempts to answer the question "What is the best way to structure a grammar?" and in particular to investigate "the relation between the sequence of grammars that the child adopts, and the basic formation of the grammar." One of the answers of [31] is very similar to our approach here; namely, that "the grammar is arranged along the lines of *sub-grammars...* so that the child passes from one to the next." However, [31] suffers from two major shortcomings. First, this work is not informed by actual data: the grammar fragments it provides are supposed to account for competence rather than performance. In this respect, it is interesting to quote [11, p. 56–58], who says:

> "not many people know how much can be milked from mere performance in the case of small children... [This work] is about knowledge; knowledge concerning grammar and the meanings coded by grammar. Knowledge inferred, of course, from performance..."

Second, [31] is, to use the terminology of [56, p. 136], "described in linguistically specific terms such that it is very difficult to relate them to cognition in other psychological domains"; in other words, this work is only accessible to experts in contemporary generative approaches to syntax.

These two issues are remedied in [61]. Here, the Universal Grammar is implemented as a Unification-Based Generalised Categorial Grammar, embedded in a default inheritance network of lexical types. The learning algorithm receives input from a corpus of spontaneous child-directed transcribed speech, annotated with logical forms, and sets the parameters based on this input. This framework is used as a basis for investigating several aspects of language acquisition from data, focusing on the acquisition of subcategorization frames and word order information. However, this work is still embedded in the principles and parameters framework, assuming a (very elaborate) universal grammar which is expressed in unification-based categorial grammar and a large set of parameters, including 89 categorial parameters, 18 word order parameters etc. As the work focuses on the acquisition of verb arguments, only 1517 carefully selected sentences were selected and annotated (out of a much larger corpus). Most importantly, this work concentrates on the "how" of language acquisition, discussing an architecture for learning the lexicon and the parameters of a grammar. The focus of our work, in contrast, is exactly on the component which [61] assumes given, namely the internal representation of the grammar, and how this grammar changes as language develops. Our work is not committed to the full details of Government-Binding theory (or any other linguistic theory, for that matter); rather, we expect child grammars to naturally emerge from actual data rather than comply with predefined principles. For a detailed, well-argued criticism of the nativist approach see [54].

Building on the work of [61], but assuming very limited innate linguistic mechanisms (e.g., the ability to recognize objects, segment words or combine constituents), and not subscribing to the principles and parameters discipline, [12] proposes a computational model for first language acquisition. The model, which infers linguistic structure from data, is statistical and can cope with noisy input. The model is shown to outperform TLA on simple learning tasks. Again, this works focuses on the "how" rather than the "what".

Surprisingly, very few works address the issue of accounting for child language data by means of formal grammars. It is worth mentioning that [11] himself developed formal grammars for describing the utterances in his corpus. He says (p. 56):

> "...the requirement to be fully explicit and develop rules that will really derive the sentences you have obtained forces a kind of intense examination and continuing re-examination of the data, and so is a good way to get to know it very well...."

We do not stop at writing the grammars, and do not limit ourselves to the goal of getting to know the data. Rather, we investigate the possible changes in grammar as it develops and the constraints imposed on such changes. Perhaps the closest approach to the one we propose here is [25], which sketches

> "...a solution to the string-to-structure problem in first language acquisition within a set of emergentist assumptions that minimizes the need to assume innate linguistic knowledge, minimizes demands for linguistic analysis by the language learner, and exploits the projection of lexical properties of words. These conceptual constraints minimize the number, complexity, and diversity of principles that have to develop in learning a grammar."

[25] sketches, very briefly, some milestones in language acquisition, and models them by means of an HPSG grammar which is gradually being developed to account for more data. The two most important principles for [25] are assuming no innate linguistic knowledge (rather, she assumes general cognitive capabilities such as "the ability to discriminate kinds of things", "the capacity to add to a store of propositional knowledge", or "the ability to identify what is missing when something is missing"); and proposing a *monotonic* theory of grammar development, although exactly what is meant by monotonicity is not explicitly defined.

[25] advocates the use of HPSG as the framework in which to model language acquisition. HPSG is chosen to demonstrate the emergentist theory mainly because of its underlying architecture, in particular the use of a type signature and, separately, a set of well-defined formal principles. The signature provides an ontology of entities (e.g., linguistic signs), and the principles constrain and restrict those entities. [25] shows that it is possible to account for several stages in child language development by means of a single mechanism: "incremental and largely monotonic changes to a type hierarchy that constitutes an increasingly less skeletal constraint-based grammar".

The main drawbacks of this work are two. First, while [25] lists several fragments of the proposed HPSG grammar, those fragments are never tested and their predictive power is never contrasted with actual corpora recording child language use. Second, the possible changes that grammars may undergo in the process of language development are never spelled out. Two of our goals in this work are to remedy these problems, by providing actual grammars that can (and are) tested against the CHILDES database corpora; and by formally defining a *grammar refinement* operator that will account for the possible changes in grammar development.

2   Research Objectives

This work can be divided into two main sub-tasks: developing grammars for representative stages of child language data; and developing grammar refinement operators that will explain the possible changes in grammars as language develops.

2.1   Grammar Development

Our departure point is the item-based, functional, emergentist theory of grammar acquisition; our grammars are therefore highly lexicalized, and assume few innate components. However, we do expect certain universal patterns to emerge from the multilingual data that we explore, and these patterns indeed inform and guide the development of the grammars. Faithful to the item-based model of language acquisition, the grammars that we develop are highly lexicalized (typed) unification grammars. As pointed out above, such grammars provide means for expressing deep linguistic structures in an integrated way. We use the LKB system [19] as our grammar development environment.

In the first phase we manually developed a few grammars for some representative corpora, focusing on the *Eve* corpus [11,39] and the *Seth* corpus [44,45]. In the next phase we will extend the coverage of the grammars to more corpora, longitudinally accounting for a small number of individuals along the first few years of language acquisition. We will consequently extend the coverage of the grammars to cover also data in other languages, for similar language development stages.

The grammars are all applied to the data in the corpus, producing deep linguistic structures describing the syntax, but also some aspects of the morphology and the semantics, of the utterances observed in the data. In particular, it is possible to automatically and deterministically convert the structural descriptions produced by the grammars to the functional annotation of [51].

2.2   Grammar Refinement

A major research question that we leave for future research involves the possible changes that a grammar can undergo as language develops; we use the term *grammar refinement* to refer in general to any of several operators which may

convert one grammar, accounting for some layer of language command, to a subsequent grammar, accounting for more developed linguistic capacity. Grammar refinement must have access to all the components of a grammar, and in particular to the type signature on which a grammar is based and to the lexicon.

Current work in unification grammars focuses on issues of grammar engineering. In a recent work, [18] define the concept of signature *modules* for unification grammars and show how several modules can combine through a *merge* operation. They introduce the concept of *partially specified signature (PSSs)* to abstract away from specific type- and feature-names in grammars. This work is extended in [55], where PSSs are extended to modules, facilitating the combination of grammar fragments, which interact via an abstract interface, and which can be combined as needed to form full grammars. While this work is preliminary, it sets the stage for several possible developments in collaborative grammar engineering. It is also a very promising starting point for defining "core" grammars which embody universal principles, and which can then be combined with more language specific and parochial modules and constraints.

We believe that these results can provide the infra-structure for defining grammar refinement operators. Operations such as splitting an existing type to several subtypes (reflecting the child's ability to make finer distinctions), or adding appropriate features to types (reflecting a more complex representation of a given sign), are easily definable with PSSs. Lexical acquisition can be accounted for by means of *grammar union* [63,64]. Acquisition and refinement of constructions can also be defined as changes to the type signature, following the encoding of constructions advocated by, e.g., [50]. In sum, unification grammars in general and those that are based on typed feature structures in particular provide suitable means for modeling language development; we hope to make significant contributions to such modeling in future extensions of this work.

3   Formal Grammars of Children's Language

In the preparation of the successive series of grammars for a single child's language, we are constrained by three requirements. First, we are bound to be faithful to the data: the grammars must adequately account for the kind of utterances observed in the CHILDES data that drive this research, and the structures they induce on utterances must be compatible with the existing syntactic annotation of [51] (henceforth, GRASP annotation). Second, our grammars must be consistent with existing psycholinguistic theories which ground the formal description in well-founded experimental and theoretical results. Formally speaking, it is always possible to provide infinitely many different grammars for any given finite language; it is the predictions of the grammar and their theoretical adequacy which distinguish a good grammar from a better one. Finally, the rules that we stipulate must be *learnable*: one can imagine a postulated rule that is generatively adequate and psycholinguistically plausible, but unlearnable. We will therefore strive to define grammar refinement operators that are compatible with existing theories of learnability.

The forces that drive grammar development are manifold; a good grammar must simultaneously account for and generate phonological, intonational, morphological, syntactic, semantic and pragmatic structures. Since this work is preliminary, we are unable to provide such wealth of information, although the framework in which we work certainly facilitates it. Two examples suffice to delineate the types of linguistic structure that our grammars abstract over.

It is clear that intonation plays a major part in (some) children's first utterances [44]. The early speech of some children is interspersed with *fillers*, and "One function underlying production of these early fillers seems to be preservation of the number of syllables in and/or the prosodic rhythm of the target" [46, p. 234]. Along the same lines, [20] proposes that "much of the null and variable occurrence of functional categories in early speech can be more accurately explained by appealing to *phonological* rather than *syntactic* aspects of children's developing grammars". However, since a formal theory of phonological and intonational structures and their interaction with other components of the grammar, even in adult language, is still unavailable, we suppress a discussion of these aspects in the sequel.

Similarly, [42] demonstrate that the very first utterances of children are better understood from a functional (i.e., pragmatic) point of view: "children's earliest words express specific communicative function" that a focus on syntax and semantics fails to analyze correctly. We believe that information structure, speech-act and pragmatics in general are essential to understanding child language, but our current grammars do not specify such information, which is not sufficiently well understood to be formalized.

3.1   Preliminaries

LKB grammars consist of three main components: a signature, specifying the type hierarchy along with constraints on types; grammar rules, which specify phrase structure combinatorics; and a lexicon, associating feature structures with words. The examples below use the LKB syntax, which is described in detail in [19].

We begin with an inventory of general-purpose types, such as lists and difference lists (figure 1). These are needed mostly for technical reasons at this preliminary stage of the grammar, so we only mention them in passing.

Next, we associate words with *part of speech (POS)* categories. It is clear that children can distinguish among some different POS categories from a very early age [30]. Several works address the question of the acquisition of POS categories, and the psychological, neurological and computational underpinnings of this process have been extensively investigated and described [33,34,35]. A classification of POS categories is manifested in our grammar as in figure 2; *co* is a communicator, e.g., *hi, bye, okay, please*, and *fil* is a filler, e.g., *uh*.

Similarly, we assume a set of *grammatical relations (GRs)* of the same inventory that is used by GRASP [51] to annotate the corpora. In this early stage of the grammar, we cannot assume that the child had learned any of those, so we resort to the general specification:

```
string := *top*.
*list* := *top*.
*ne-list* := *list* & [ FIRST *top*, REST *list* ].
*null* := *list*.
*diff-list* := *top* & [ LIST *list*, LAST *list* ].
*empty-diff-list* := *diff-list* & [ LIST #list, LAST #list ].
```

Fig. 1. General purpose types

```
cat := *top*.
  adj := cat.
  adv := cat.
  co := cat.
  fil := cat.
  nominal := cat.
    n := nominal.
    prop := nominal.
  prep := cat.
  v := cat.
```

Fig. 2. Parts of speech

```
gr := *top*.
```

Furthermore, we define a structure for dependency relations. A dependency is a triple consisting of a type, which is a GR, and a head and a dependent, which are lexical structures:

```
dep := *top* &
[ GRTYPE gr,
  GRDEP lex,
  GRHEAD lex ].
```

Lexical structures consist of a string, representing the standard orthography of the lexical item, and a GR. The string is merely a convenient abstraction over the actual phonological structure of the word, which should be properly modeled as an extension of the present work. The reason for the GR is that oftentimes it is the lexical item which determines the grammatical relation of some construction; we return to this below.

```
lex := *top* &
[ ORTH string,
  DEP gr ].
```

Syntactic structures consist of a part of speech CATegory; a sub-category, which we take to be a list of complements; and a list of specifiers, as in HPSG. The SPR list will be used by non-heads to specify the heads they modify; it will

either have a single element or be empty. In contrast, SUBCAT lists are used by heads to specify their complements and may have zero or more elements.

```
syn := *top* &
[ CAT cat,
  SPR *list*,
  SUBCAT *list* ].
```

Finally, we define *constructions*. A construction [23,24,57] is a collection of linguistic information pertaining to the form, combinatorics and meaning of an utterance, similarly to HPSG's *signs*. In our grammars, we roughly approximate constructions by specifying the lexical and syntactic information structures of words and phrases, as well as a list of grammatical relations GRS in lieu of semantics. Additionally, the feature STRING lists the surface form of the construction, as in figure 3.

```
construction := *top* &
[ LEX lex,
  SYN syn,
  GRS *diff-list*,
  STRING *diff-list* ].
```

Fig. 3. Constructions

Of course, figure 3 only lists the very basic constructions, and sub-types of the type *construction* are defined and refined as the grammar develops. For our first grammar we can assume that two sub-types are defined, namely *lexeme* and *phrase*; the former is associated with lexical items, whereas the latter has internal structure, expressed via (at least) the feature DTR1, standing for 'daughter 1', as in figure 4.

```
lexeme := construction &
[ LEX #lex,
  GRS *empty-diff-list*,
  STRING [ LIST [ FIRST #lex , REST #last ], LAST #last ] ].

phrase := construction &
[ DTR1 construction ].
```

Fig. 4. Lexemes and phrases

3.2   Earliest Language

The first grammar reflects the very first utterances recorded in the corpus; these are usually one-word utterances and *holophrases* [59, p. 261]. We constructed a lexicon of all such items which occur in the corpus; each item is associated with a

```
Julie   :=  [ LEX [ ORTH "Julie" ] ] & lex_prop .
back    :=  [ LEX [ ORTH "back" ] ] & lex_adv .
byebye  :=  [ LEX [ ORTH "byebye" ] ] & lex_co .
cake    :=  [ LEX [ ORTH "cake" ] ] & lex_n .
came    :=  [ LEX [ ORTH "came" ] ] & lex_v .
```

Fig. 5. Some lexical items

```
lex_adj := lexeme & [ SYN [ CAT adj ] ].
lex_adv := lexeme & [ SYN [ CAT adv ], LEX [ DEP jct-rel ] ].
lex_co := lexeme & [ SYN [ CAT co ] ].
lex_fil := lexeme & [ SYN [ CAT fil ] ].
lex_n := lexeme & [ SYN [ CAT n ] ].
lex_prop := lexeme & [ SYN [ CAT prop ] ].
lex_prep := lexeme & [ SYN [ CAT prep ] ].
lex_v := lexeme & [ SYN [ CAT v ] ].
```

Fig. 6. Sub-types of *lexeme*

feature structure of type *lexeme*, reflecting at least its orthography and syntactic (POS) category. Some examples are listed in figure 5. The types *lex_prop, lex_adv* etc. are defined as in figure 6.

These definitions are sufficient for handling the one-word utterances that are observed in the first files of the Seth corpus.

3.3   The Emergence of Constructions

The next stage of language development allows for combinations of words and holophrases. For example, the first multi-word utterances in the Seth corpus are *came in; come in; had cake; had chocolate; orange juice*; and *right back*. We view the latter two as holophrases, but the first four, which are augmented in subsequent data by utterances such as *Dabee toast* and *swing high*, justify the introduction of item-based constructions to the grammar.

Exactly how constructions are learned by children remains largely unknown [32], and we do not attempt to address this question directly here. However, at least three types of constructions must be supported by a formal account of grammar development. First, basic items can be concatenated, yielding "successive single-word utterances" [6]; according to [8], these are the first indications of syntax in speech. Second, *pivot-schemas* allow the combination of one highly frequent "event-word" (e.g., *more, It's, I*) with a variety of items. Third, *item-based constructions* allow more complex and more abstract combination, in which the construction itself adds, for the first time, to the semantics of the full utterance [59, pp. 263-264].

Let us focus first on utterances such as *came in; come in* and *had cake; had chocolate*. By the end of the language learning process we may want to say that such utterances reflect *verb–adjunct* and *verb–object* relations, respectively. This

is indeed how these utterances are annotated in the corpus. Therefore, during the learning process the child must somehow acquire these constructions. A possible way of stepping through this process is outlined below.

First, the grammar must allow binary combinations of words; this is captured by the type *binary-phrase*, which is a sub-type of *phrase*. A *binary-phrase* has two daughters, DTR1 and DTR2. Furthermore, the list of GRS of the mother is obtained by concatenating the same lists in the two daughters, and the same holds for STRING. In order to distinguish between the two daughters, we introduce the type *headed-phrase*, in which an additional feature, HEAD-DTR, reflects (and is reentrant with) the head daughter, which can be either DTR1 or DTR2. The head daughter shares its LEX structure and its main CATegory, but not necessarily its SUBCATegory, with its mother. We next define a *binary-headed-phrase*, a sub-type of both *binary-phrase* and *headed-phrase*, with an additional feature, NON-HEAD-DTR. These definitions are given in figure 7.

```
binary-phrase := phrase &
[ DTR1 [ GRS [ LIST #first, LAST #middle ],
         STRING [ LIST #str1, LAST #str3 ] ],
  DTR2 [ GRS [ LIST #middle, LAST #last ],
         STRING [ LIST #str3, LAST #str2 ] ],
  GRS [ LIST [ FIRST dep, REST #first ], LAST #last ],
  STRING [ LIST #str1, LAST #str2 ] ].

headed-phrase := phrase &
[ LEX #lex,
  SYN [ CAT #cat ],
  HEAD-DTR construction & [ LEX #lex, SYN [ CAT #cat ] ] ].

binary-headed-phrase := headed-phrase & binary-phrase &
[ NON-HEAD-DTR construction ].
```

Fig. 7. Binary phrases

In both verb–adjunct and verb-object constructions, the head is the first daughter; furthermore, in both the second daughter is a dependent of the first. We therefore define two sub-types of *binary-phrase*, namely *head-first* and *dep-second*, which reflect these observations. A major difference between adjuncts and objects is that the latter are subcategorized for by the head verb; we define two subtypes of *binary-phrase* which account to this distinction: in *subcat-unchanged*, the SUBCATegorization of the mother and the head daughter are identical, whereas in *subcat1* constructions the head daughter specifies its object argument, which is removed from the SUBCATegorization list that is propagated to the mother. With these definitions in place, *obj-construction* and *jct-construction* are defined as in figure 8.

The only grammar rules that are needed for this elementary stage are two: one for each type of constructions. These rules are depicted in figure 9. While the

```
head-first := binary-headed-phrase &
[ HEAD-DTR #head-dtr,
  DTR1 #head-dtr,
  NON-HEAD-DTR #non-head-dtr,
  DTR2 #non-head-dtr ].

dep-second := binary-phrase &
[ GRS [ LIST [ FIRST [ GRHEAD #dep1, GRDEP #dep2 ] ] ],
  DTR1 [ LEX #dep1 ],
  DTR2 [ LEX #dep2 ] ].

subcat-unchanged := binary-headed-phrase &
[ SYN [ SUBCAT #subcat, SPR #spr ],
  HEAD-DTR [ SYN [ SUBCAT #subcat, SPR #spr ] ] ].

subcat1 := binary-headed-phrase &
[ SYN [ SUBCAT #rest, SPR #spr ],
  HEAD-DTR [ SYN [ SPR #spr ] ],
  DTR2 #head-dtr,
  DTR1 [ SYN [ SUBCAT [ FIRST #head-dtr, REST #rest ] ] ] ].

obj-construction := head-first & dep-second & subcat1 &
  [ GRS.LIST.FIRST.GRTYPE obj-rel ].
jct-construction := head-first & dep-second & subcat-unchanged &
  [ GRS.LIST.FIRST.GRTYPE jct-rel ].
```

Fig. 8. Types of binary phrases

rules induce isomorphic trees on the utterances *had cake* and *come in* (figure 10), the feature structures induces by the grammar on the two strings are significantly different (figure 11). In fact, words such as *in* are categorized as *adv*erbs for convenience only. We do not claim that the child had acquired the category of adverbs yet; rather, it is clear that he uses some verbs with some adjuncts, as in *come in* or *swing high*. It is also clear that these constructions are different from, say, *had cake* or *had chocolate*. We therefore classify "words which can follow verbs" as adjuncts, and "nouns which can follow the verb *have*" as objects. Also, the rule which allows objects to combine with verbs is general: it can apply to any verb. It is well known that only few verbs are used in early language, and that argument structure is acquired in an item-based fashion [24]. The rule could therefore have been stated in terms of a particular verb, by specifically referring to the LEX feature of the verb, or to a group of verbs, by specifying a sub-type of *verb*.

The first utterance in the Seth corpus which is clearly not an instance of either the verb–object or the verb-adjunct constructions is *Dabee toast* (Dabee is the child's grandmother). We treat such utterances as *vocative* constructions: the first word of the construction is a proper name, and the second is unconstrained. Few additions to the grammar are required in order to account for such

```
vp_v_np_rule := obj-construction &
[ SYN [ CAT v ],
  DTR2 [ SYN np ] ] .

vp_v_advp_rule := jct-construction &
[ SYN [ CAT v ],
  DTR2 [ SYN [ CAT adv ] ] ] .
```

Fig. 9. Two phrase-structure rules

Fig. 10. Two trees

utterances. In a vocative construction, the head is the second daughter, and the first word is a dependent of the second; we therefore add the types *head-second* and *dep-first*, analogously to *head-first* and *dep-second*. Also, the resulting phrase is saturated, as no further modification is allowed. We subsequently add a *sub-cat0* construction which specifies that the SUBCATegorization list of the mother is empty. A *voc-construction* is a sub-type of both. One rule is added, too, as shown in figure 12.

3.4   Towards Full Syntax

The infra-structure set up above facilitates a formal account of the various types of constructions learned at this early stage. In the first files of the Seth corpus different kinds of multi-word utterances are observed, including *Dabee toast; swing high; Uncle Bob; bird outside; I bet; kiss it; come see me; put shoes on; take a bite* and *who loves you*. These phrases represent various grammatical relations, and while some of them are pivot-based or item-based, others may be the beginning of the emerging syntax. We describe below the extensions to the grammar which are required to account for these new data.

Let us first account for the emergence of subjects, as in *I bet*. We add a sub-type of *nominal*, called *pro*noun, of which *I* is an instance. We also add a sub-type of *gr*, namely *subj-gr*. A subject construction is a special case of the newly introduced *specifier–head* construction, in which the first daughter is the specifier of the head of the second daughter, and SUBCAT is empty and shared between the mother and the head daughter. Also, the first constituent depends on the second, hence the type specification and grammar rule listed in figure 13.
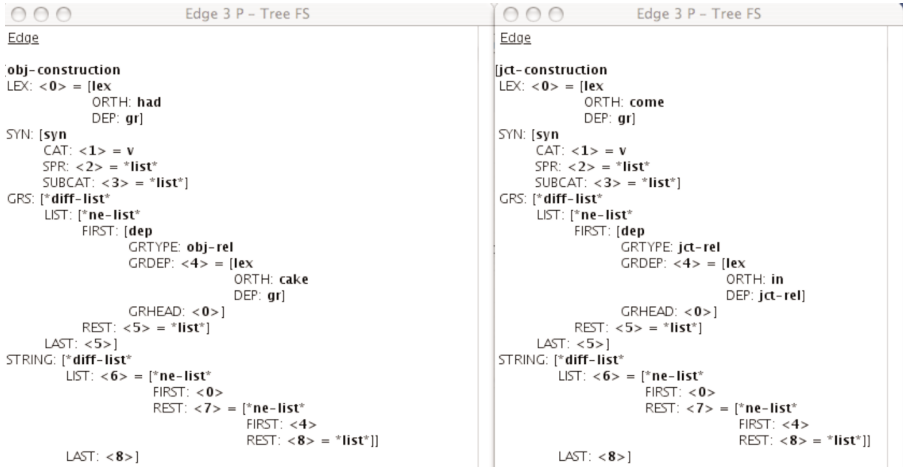
Fig. 11. Feature structures of two two-word utterances

```
head-second := binary-headed-phrase &
[ HEAD-DTR #head-dtr,
  DTR2 #head-dtr,
  NON-HEAD-DTR #non-head-dtr,
  DTR1 #non-head-dtr ].

dep-first := binary-phrase &
[ GRS [ LIST [ FIRST [ GRHEAD #dep2, GRDEP #dep1 ] ] ] ],
  DTR1 [ LEX #dep1 ],
  DTR2 [ LEX #dep2 ] ].

subcat0 := binary-phrase &
[ SYN [ SUBCAT *null*, SPR *null* ] ].

 voc-construction := dep-first & subcat0 & subcat1 &
 [ GRS.LIST.FIRST.GRTYPE voc-rel ].

vocative_rule := voc-construction &
[ DTR1 [ SYN [ CAT prop ] ] ] .
```

Fig. 12. Accounting for vocative constructions

We now move on to the emergence of determiners, as in *take a bite*. We need an additional sub-type of *cat*, namely *det*erminer, and an additional sub-type of *gr*, namely *det-rel*. Like subjects, determiner are specifiers of their heads, but unlike subjects they specify the category of their heads in their lexical types, as shown in figure 14.

```
spr-head := binary-headed-phrase & dep-first &
[ SYN [ SUBCAT #subcat & *null*, SPR *null* ],
  HEAD-DTR #head-dtr & [ SYN [ SUBCAT #subcat ] ],
  DTR1 #spr-dtr,
  DTR2 #head-dtr &  [ SYN [ SPR [ FIRST #spr-dtr ] ] ] ].

subj-construction := spr-head & [ GRS.LIST.FIRST.GRTYPE subj-rel ].

s_np_vp_rule := subj-construction &
[ SYN clause,
  DTR1 [ SYN np ],
  DTR2 [ SYN [ CAT v ] ] ] .
```

Fig. 13. Accounting for subject constructions

```
det-construction := spr-head & [ GRS.LIST.FIRST.GRTYPE det-rel ].

lex_det := lexeme &
           [ SYN [ CAT det, SPR [ FIRST [ SYN np ], REST *null* ] ] ].

np_det_n_rule := det-construction &
[ SYN np,
  DTR1 lex_det,
  DTR2 lex_n ] .
```

Fig. 14. Accounting for determiner constructions

The trees induced by the grammar on *I bet* and *take a bite* are depicted in figure 15. Of course, with the available constructions the grammar can generate more complex structures, such as *I take a bite*, which are not observed in the data at this early stage. We take this to be a performance constraint on grammatical competence, which is better realized by limiting, e.g., the length of the utterances or the degree of tree nesting.

Immediately after the appearance of determiners, quantifiers are observed in the data in very similar constructions: *some toast, some juice*. In addition to the introduction of the types *qn* (a sub-type of *cat*) and *quant-rel* (a sub-type of *gr*), no additional construction type is needed. Rather, we specify quantifiers lexically analogously to determiners, and add a designated rule which is almost identical to the determiner rule, see figure 16.

Another construction that is observed in this early stage involves *communicators*: words such as *hi, oh, please* or *yes*, followed by some other short utterance, as in *oh good* or *please Daddy*. These are accounted for by the introduction of a new sub-type of *gr*, namely *com-rel*; a new construction type, *com-construction*; and an additional rule, depicted in figure 17.

At a very early age children begin to develop more complex syntactic structures. Seth, for example, produces an interrogative construction at the age of
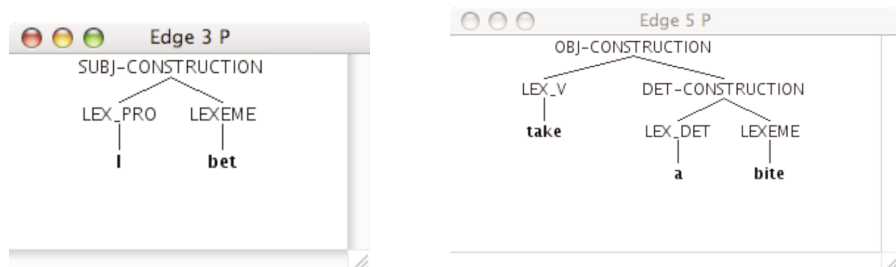
Fig. 15. Two more trees

```
lex_qn :=  lexeme &
           [ SYN [ CAT qn,  SPR [ FIRST [ SYN np ], REST *null* ] ] ].

np_qn_n_rule := det-construction &
[ SYN np,
  DTR1 lex_qn,
  DTR2 lex_n ] .
```

Fig. 16. Accounting for quantifiers

```
com-construction := head-second & dep-first & subcat1 &
                    [ GRS.LIST.FIRST.GRTYPE com-rel ].

communicator_rule := com-construction &
[ DTR1 lex_co
] .
```

Fig. 17. Accounting for communicator constructions

twenty months. It is *what geese say*, whose meaning can be interpreted from the following utterance in the corpus: his father repeating *what do the geese say?* In what follows we provide an account of such constructions, inspired by the HPSG treatment of "movement", or transformations, using *slash* categories.

First, we split the type *pronoun* into two sub-types, *pers-pro* and *wh-pro*, distinguishing between personal and interrogative pronouns, respectively. We similarly split *lex-pro* into *lex-pers-pro* and *lex-wh-pro*. Also, in order to limit the extraction of the object in utterances like *what geese say* to transitive verbs only, we define the type *lex-trans*, which is a sub-type of *lex-v* which has a single, noun phrase element on its SUBCAT list.

Next, we augment the type *construction* by adding a feature SLASH, whose value is a list. We expect at most one element on slash lists, an instance of the type *slash*:

```
slash := *top* &
[ CONST construction,
  SLASH-DEP dep ].
```

When an expected argument is "moved", as in the case of the object of *say* in the running example, it is removed from the SUBCAT list of its mother and stored in the SLASH list. This is accounted for by *gap* constructions. Such elements are expected to be realized higher up the tree as *wh-pro*nouns; they are then matched against the stored element in the SLASH list of the second daughter, through the *filler* construction. These two constructions, along with some additional supporting types, are depicted in figure 18 (some additional, minor modifications are required in the grammar in order to propagate the value of the SLASH feature along the derivation). The resulting tree is depicted in figure 19.

4   Evaluation

In order to evaluate the plausibility of the grammars, we extracted from the structures that they induce on child utterances the sequence of grammatical relations, and compared them to the manual GRASP annotations. We briefly describe the methodology and discuss the results of the evaluation in this section.

```
unary-phrase := phrase &
[ DTR1 [ GRS #grs, STRING #str ],
  GRS #grs,
  STRING #str ].

unary-headed-phrase := headed-phrase & unary-phrase &
[ HEAD-DTR #dtr,
  DTR1 #dtr ].

gap := unary-headed-phrase &
[ SYN [ SUBCAT #rest, SPR #spr ],
  SLASH [ FIRST [ CONST #missing-dtr,
                  SLASH-DEP [ GRTYPE obj-rel ] ] ],
  DTR1 [ SYN [ SUBCAT [ FIRST #missing-dtr, REST #rest ], SPR #spr ] ] ].

filler := head-second & dep-first &
[ DTR2 [ SYN #syn, SLASH [ FIRST [ CONST #dtr,
                                   SLASH-DEP [ GRTYPE #grtype ] ] ] ],
  DTR1 #dtr & [ SYN [ CAT wh-pro ] ],
  SYN #syn,
  GRS.LIST.FIRST.GRTYPE #grtype ].
```
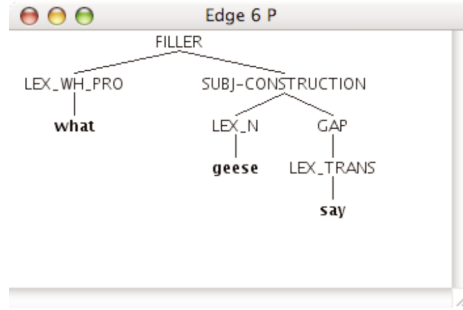
Fig. 18. Gaps and fillers

Fig. 19. Filler-gap constructions

First, we manually annotated a subset of the Seth corpus, focusing only on child utterances and ignoring any adult ones, using the GRASP scheme. We then extracted from this subset only those utterances that are *well-formed*, i.e., include only words which occur in the lexicon (as opposed to filler syllables, un-recognized sounds etc.) These 516 utterances constitute our development corpus. We then parsed the utterances with the set of grammars described above. From the resulting feature structures we extracted the values of the GRS feature, and converted them to GRASP format.

Table 1 lists the number of errors of the parser, in terms of unlabeled and labeled dependencies, with respect to the manually annotated corpus. Since the corpus contains many repetitions we present both the number of utterance tokens (total) and the number of utterance types (unique occurrences).

Table 1. Evaluation results, development set

File	# utterances	# of GRs	Grammar Errors			
			unlabeled		labeled	
			#	%	#	%
0a tokens	107	234	48	21	48	21
0a types	38	82	6	7	6	7
1a tokens	108	223	3	1	7	3
1a types	38	79	3	4	5	6
2a tokens	115	247	16	6	18	7
2a types	45	105	13	12	15	14
2b tokens	68	142	12	8	12	8
2b types	30	64	6	9	6	9
3a tokens	118	240	0	0	0	0
3a types	34	72	0	0	0	0
Total tokens	516	1086	79	7.9	85	7.8
Total types	185	402	28	6.9	32	7.9

5   Conclusion

We presented a sequence of grammars which adequately cover the first stages of the emergence of syntax in the language of one child. The structures produced by the grammars were evaluated against a manually annotated corpus, revealing a low error rate (below 10%). These are preliminary results, and the evaluation is very basic, but we believe that the results are promising.

Space considerations prevent us from listing the complete grammars. However, we note that the changes introduced in each grammar, compared to the previous one, consist only of two operations: adding types, by splitting a single type to two or more sub-types; and adding constraints on types, in the form of additional features or reentrancy constraints. This is an encouraging outcome: in a very well-founded, mathematical sense, our grammars are monotonically increasing.

This preliminary work will be continued in three main tracks. First, we intend to continue the development of the grammars, accounting for more constructions and evaluating the accuracy on the entire Seth corpus, as well as on other manually annotated child language corpora. Second, we will develop mechanisms of grammar engineering that will facilitate both the definition of refinement operators and the modular development of sequences of grammars. Finally, we will account for competition among constraints, setting the stage for a theory which could also explain *how* these grammars can be learned from data.

Acknowledgements

References

1. Abney, S.P.: Stochastic attribute-value grammars. Computational Linguistics 23(4), 597–618 (1997)
2. Bates, E., MacWhinney, B.: Competition, variation, and language learning. In: MacWhinney, B. (ed.) Mechanisms of language acquisition, ch. 6, pp. 157–193. Lawrence Erlbaum Associates, Hillsdale (1987)
3. Bender, E.M., Flickinger, D., Fouvry, F., Siegel, M.: Shared representation in multilingual grammar engineering. Research on Language and Computation 3, 131–138 (2005)
4. Berman, R.A.: Between emergence and mastery: The long developmental route of language acquisition. In: Berman, R.A. (ed.) Language development across childhood and adolescence. Trends in Language Acquisition Research, vol. 3, pp. 9–34. John Benjamins, Amsterdam (2004)
5. Bertolo, S.: Acquisition, formal theories of. In: Wilson, R.A., Keil, F.C. (eds.) The MIT Encyclopedia of the Cognitive Sciences. Bradford Books (September 2001)

6. Bloom, L.: One word at a time: The use of single word utterances before syntax. Mouton, The Hague (1973)
7. Borensztajn, G., Zuidema, W., Bod, R.: Children's grammars grow more abstract with age - evidence from an automatic procedure for identifying the productive units of language. In: Proceedings of CogSci (2008)
8. Branigan, G.: Some reasons why successive single word utterances are not. Journal of Child Language 6, 411–421 (1979)
9. Brent, M.R.: Advances in the computational study of language acquisition. In: Brent, M.R. (ed.) Computational Approaches to Language Acquisition, Cognition Special Issues, ch. 1, pp. 1–38. The MIT Press, Cambridge (1997)
10. Brent, M.R. (ed.): Computational Approaches to Language Acquisition. Cognition Special Issues. The MIT Press, Cambridge (1997)
11. Brown, R.: A first language: the Early stages. Harvard University Press, Cambridge (1973)
12. Buttery, P.J.: Computational models for first language acquisition. Technical Report UCAM-CL-TR-675, University of Cambridge, Computer Laboratory (November 2006)
13. Carpenter, B.: The Logic of Typed Feature Structures. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, Cambridge (1992)
14. Chomsky, N.: Aspects of the theory of syntax. MIT Press, Cambridge (1965)
15. Chomsky, N.: Language and Mind. Harcourt Brace Juvanovich, New York (1968)
16. Chomsky, N.: Rules and representations. Behavioral and Brain Sciences 3, 1–61 (1980)
17. Chomsky, N., Lasnik, H.: The theory of principles and parameters. In: Jacobs, J., von Stechow, A., Sternefeld, W., Vannemann, T. (eds.) Syntax: An International Handbook of Contemporary Research. Walter de Gruyter, Berlin, pp. 506–569 (1993); Reprinted in Chomsky, N.: The Minimalist Program, pp. 13–127. MIT Press, Cambridge (1995)
18. Cohen-Sygal, Y., Wintner, S.: Partially specified signatures: a vehicle for grammar modularity. In: Proceedings of Coling–ACL 2006, Sydney, Australia, July 2006, pp. 145–152 (2006)
19. Copestake, A.: Implementing Typed Feature Structure Grammars. CSLI Publications, Stanford (2002)
20. Demuth, K.: On the 'underspecification' of functional categories in early grammars. In: Lust, B., Suñer, M., Whitman, J. (eds.) Syntactic theory and first language acquisition: cross-linguistic perspectives, pp. 119–134. Lawrence Erlbaum Associates, Hillsdale (1994)
21. Gibson, E., Wexler, K.: Triggers. Linguistic Inquiry 25(4), 407–454 (1994)
22. Mark Gold, E.: Language identification in the limit. Information and Control 10(5), 447–474 (1967)
23. Goldberg, A.: Constructions. A Construction Grammar approach to argument structure. University of Chicago Press, Chicago (1995)
24. Goldberg, A.: Constructions at Work: the nature of generalization in language. Oxford University Press, Oxford (2006)
25. Green, G.M.: Modelling grammar growth; universal grammar without innate principles or parameters, Unpubl. ms., University of Illinois (July 2003)
26. Hinrichs, E.W., Meurers, W.D., Wintner, S.: Linguistic theory and grammar implementation. Research on Language and Computation 2, 155–163 (2004)
27. Johnson, M., Riezler, S.: Statistical models of language learning and use. Cognitive Science 26(3), 239–253 (2002)

28. Kaplan, R., Bresnan, J.: Lexical functional grammar: A formal system for grammatical representation. In: Bresnan, J. (ed.) The Mental Representation of Grammatical Relations, pp. 173–281. MIT Press, Cambridge (1982)
29. King, T.H., Forst, M., Kuhn, J., Butt, M.: The feature space in parallel grammar writing. Research on Language and Computation 3, 139–163 (2005)
30. Labelle, M.: The acquisition of grammatical categories: a state of the art. In: Cohen, H., Lefebvre, C. (eds.) Handbook of categorization in cognitive science, pp. 433–457. Elsevier, Amsterdam (2005)
31. Lebeaux, D.: Language acquisition and the form of the grammar. John Benjamins, Amsterdam (2000)
32. Levy, Y., Schlesinger, I.M.: The child's early categories: approaches to language acquisition theory. In: Levy, et al. (eds.) [33], ch. 9, pp. 261–276.
33. Levy, Y., Schlesinger, I.M., Braine, M.D.S. (eds.): Categories and Processes in Language Acquisition. Lawrence Erlbaum Associates, Hillsdale (1988)
34. Li, P.: Language acquisition in a self-organizing neural network model. In: Quinlan, P. (ed.) Connectionist models of development: Developmental processes in real and artificial neural networks. Psychology Press, Hove (2003)
35. Li, P., Farkas, I., MacWhinney, B.: Early lexical development in a self-organizing neural network. Neural Network 17, 1345–1362 (2004)
36. MacWhinney, B.: Rules, rote, and analogy in morphological formations by Hungarian children. Journal of Child Language 2, 65–77 (1975)
37. MacWhinney, B.: Models of the emergence of language. Annual Review of Psychology 49, 199–227 (1998)
38. MacWhinney, B.: The emergence of language. In: Carnegie Mellon Symposia on Cognition. Lawrence Erlbaum Associates, Mahwah (1999)
39. MacWhinney, B.: The CHILDES Project: Tools for Analyzing Talk, 3rd edn. Lawrence Erlbaum Associates, Mahwah (2000)
40. MacWhinney, B.: A multiple process solution to the logical problem of language acquisition. Journal of Child Language 31, 883–914 (2004)
41. MacWhinney, B.: A unified model of language acquisition. In: Kroll, J., De Groot, A. (eds.) Handbook of bilingualism: Psycholinguistic approaches. Oxford University Press, Oxford (2004)
42. Ninio, A., Snow, C.E.: Language acquisition through language use: the functional source of children's early utterances. In: Levy, et al. (eds.) [33], ch. 1, pp. 11–30
43. Niyogi, P., Berwick, R.C.: A language learning model for finite parameter spaces. Cognition 61(1-2), 161–193 (1996)
44. Peters, A.M.: The Units of Language Acquisition. In: Monographs in Applied Psycholinguistics. Cambridge University Press, New York (1983)
45. Peters, A.M.: Strategies in the acquisition of syntax. In: Fletcher, P., MacWhinney, B. (eds.) The handbook of child language, pp. 462–482. Blackwell, Oxford (1995)
46. Peters, A.M.: Filler syllables: what is their status in emerging grammar? Journal of Child Language 28, 229–242 (2001)
47. Pinker, S.: The Language Instinct. William Morrow and Company, New York (1994)
48. Pollard, C., Sag, I.A.: Head-Driven Phrase Structure Grammar. University of Chicago Press and CSLI Publications (1994)
49. Prince, A., Smolensky, P.: Optimality: from neural networks to universal grammar. Science 275, 1604–1610 (1997)
50. Sag, I.A.: English relative clause constructions. Journal of Linguistics 33(2), 431–484 (1997)

51. Sagae, K., Davis, E., Lavie, A., MacWhinney, B., Wintner, S.: High-accuracy annotation and parsing of CHILDES transcripts. In: Proceedings of the ACL 2007 Workshop on Cognitive Aspects of Computational Language Acquisition, Prague, Czech Republic, June 2007, pp. 25–32. Association for Computational Linguistics (2007)
52. Sagae, K., Lavie, A., MacWhinney, B.: Automatic measurement of syntactic development in child language. In: Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005), Ann Arbor, Michigan, June 2005, pp. 197–204. Association for Computational Linguistics (2005)
53. Sagae, K., MacWhinney, B., Lavie, A.: Automatic parsing of parent-child interactions. Behavior Research Methods, Instruments, and Computers 36, 113–126 (2004)
54. Sampson, G.: The 'Language Instinct' Debate, revised edn. Continuum, London (2005)
55. Sygal, Y., Wintner, S.: Type signature modules. In: de Groote, P. (ed.) Proceedings of FG 2008: The 13th conference on Formal Grammar, August, pp. 113–128 (2008)
56. Tomasello, M.: Language is not an instinct. Cognitive Development 10, 131–156 (1995)
57. Tomasello, M.: Constructing a language. Harvard University Press, Cambridge (2003)
58. Tomasello, M.: On the different origins of symbols and grammars. In: Christiansen, M.H., Kirby, S. (eds.) Language Evolution, Studies in the Evolution of Language, ch. 6, pp. 94–110. Oxford University Press, Oxford (2003)
59. Tomasello, M.: Acquiring linguistic constructions. In: Kuhn, D., Siegler, R. (eds.) Handbook of Child Psychology, pp. 255–298. Wiley, New York (2006)
60. Valiant, L.G.: A theory of the learnable. Commun. ACM 27(11), 1134–1142 (1984)
61. Villavicencio, A.: The Acquisition of a Unification-Based Generalised Categorial Grammar. Technical report UCAM-CL-TR-533, Computer Laboratory, University of Cambridge (April 2001)
62. Wexler, K., Culicover, P.W.: Formal principles of language acquisition. The MIT Press, Cambridge (1980)
63. Wintner, S.: Modular context-free grammars. Grammars 5(1), 41–63 (2002)
64. Wintner, S.: On the semantics of unification grammars. Grammars 6(2), 145–153 (2003)
65. Wintner, S.: Introduction to unification grammars. In: Ésik, Z., Martín-Vide, C., Mitrana, V. (eds.) Recent Advances in Formal Languages and Applications, ch. 13. Studies in Computational Intelligence, vol. 25, pp. 321–342. Springer, Heidelberg (2006)
66. Wintner, S.: Unification: Computational issues. In: Brown, K. (ed.) Encyclopedia of Language and Linguistics, 2nd edn., vol. 13, pp. 238–250. Elsevier, Oxford (2006)