Item-based patterns in early syntactic development

Brian MacWhinney

1. From words to combinations

Children begin language learning by producing one word at a time (Bloom 1973). It may seem obvious that children build up language by putting together small pieces into larger, more complex structures (Simon 1969). However, some researchers have argued that children cannot pick up single words from parental input without relying on additional processes such as statistical learning (Thiessen and Saffran 2007), syntactic bootstrapping (Gleitman 1990), or semantic bootstrapping (Pinker 1995, Siskind 2000). Although these processes are involved in various ways during language learning, it is not clear that they are crucially involved in word learning. Instead, as St. Augustine argued in his Confessions (1952) back in the 4th century, children pick up words because of the ways in which parents present them, often by pointing at objects directly and naming them. To explore this issue, I examined the maternal input to 16 children in the Brent Corpus in the CHILDES database (http://childes.psy.cmu.edu, MacWhinney 2000). The children in this corpus were studied between 9 and 15 months of age and the total size of the database is 496,000 words. This search shows that 23.8% of the maternal utterances are single word utterances.[1] These results indicate that Augustine's model is not that far off the mark, and that it is safe to assume that children can pick up many words without relying on additional segmentation (Aslin, Saffran, and Newport 1999) strategies and bootstrapping. Recent models of early word learning (Blanchard, Heinz, and Golinkoff 2010, Monaghan and Christiansen 2010, Rytting, Brew, and Fosler-Lussier 2010) place increasing emphasis on the role of the lexicon in guiding segmentation. Although statistical learning may help guide segmentation and thereby lexical learning, the pathway

1. The CLAN commands are:
freq +s"*MOT:" +u +re +d4 +y +.cha (locates 155906 maternal utterances)
wdlen +t*MOT +re +u *.cha (locates 37110 maternal one-word utterances)

from lexical learning to segmentation is even more central to language learning. Although this learning path may not be available for bound morphemes and function words, it is clearly available for many content words. As we will see later, learning of bound morphemes follows a similar, but slightly different path.

It is clear that child language learning is not based on the full batch recording of long sentences completely mapped to complex semantic structures, as suggested in models such as SCISSOR (Ge and Mooney 2005). Such models may seem attractive from a computational point of view, but they make implausible assumptions regarding children's memory for sentences and their contexts. They make the excessively strong assumption that children store away everything they hear, along with complete episodic encodings of the situations in which utterances occur. Despite their full control of the language, even adults do not demonstrate this level of total recall (Keenan, MacWhinney, and Mayhew 1977), and it seems still less likely that children could have this level of recall for sentences that they do not yet even comprehend. Instead, we can think of the language learning process as one in which small components are first isolated and then assembled into larger combinations, step by step. Many early forms are multimorphemic combinations that function initially as single lexical items, or what MacWhinney (1975b, 1978) called "amalgams". For example, a child who only knows the word *dishes* and not *dish*, may think that *dishes* is a mass noun with no plural suffix, much like *clothes*. In addition to these morphemic amalgams, children will pick up longer phrasal combinations as if they were single lexical items. For example, they may treat *where's the* as a single lexical item, not analyzed into three morphemes. To trace this process, we need to look closely at the actual utterances produced by children. In practice, the study of these early two- and three-word combinations has involved decades of data collection and theory construction. The Child Language Data Exchange System (CHILDES) has collected longitudinal data from hundreds of children learning dozens of languages with a particularly heavy representation of data from children producing early word combinations. The construction of this important resource was made possible by the generous contributions of hundreds of child language researchers who have made their hard-won data publicly available. Using this resource, we are now able to produce increasingly refined accounts of the process of early syntactic development.

The transition from children's first words to their first sentences is nearly imperceptible. After learning the first words, children begin to produce more and more single-word utterances. As their vocabulary grows, children begin saying words in close approximation, separated only by short pauses (Branigan 1979). For example, they may say *wanna*, followed by a short pause and then *cookie*. If the intonational contour of *wanna* is not closely integrated with that of *cookie*, adults tend to perceive this as two successive single-word utterances. However, the child may already have in mind a clear syntactic relation between the two words. As the clarity of the relations between single words strengthens, the temporal gap between the words decreases. Eventually, we judge the production of *want cookie* to be a single multi-word utterance. Across a period of several months, two- and three-word combinations such as *more cookie*, *my baby*, *hi Daddy*, and *look my birdie* become increasingly frequent. In experiments attempting to teach signed language to chimpanzees, this transition from successive single word utterances to single multiword utterances seems to occur less frequently or not at all. This has led researchers (Hauser, Chomsky, and Fitch 2002, MacWhinney 2008a, Terrace et al. 1980) to suggest that the ability to communicate using integrated combinations is uniquely supported by the human language mechanism.

2. Positional patterns

In parallel with the ongoing process of data collection, child language researchers have examined a variety of theoretical accounts of early word combinations. The goal of this work is to formulate a set of mechanisms (MacWhinney 1987) that can explain how children use the linguistic input they receive to construct productive grammatical patterns. The first attempt to provide an account of these early grammars was offered by Braine (1963, 1976). Braine suggested that early word combinations were structured in terms of "positional patterns" that involved the linear combination of two classes of words: pivot class words and open class words. Words in the pivot class could only occur in combination with open class words, whereas open class words could either appear alone, in combination with pivot class words, or in combination with other open class words. Braine referred to this system as a Pivot Grammar. His analysis of this system was backed up by experiments (Braine 1963, 1966) that showed how adults could extract word classes of this type in miniature linguistic systems (MLS). In a classic analysis, Bloom (1971) challenged the generative adequacy of the Pivot Grammar framework by emphasizing two problems. The first was the tendency for Pivot Grammar to overgenerate. For example, it would allow forms like *want take* or *my want* in which words were combined in conceptually unlikely ways. The second problem involved the analysis of open-open combinations such as *Mommy chair*. In such combinations, it is difficult to determine if the child intends “Mommy’s chair”, “Mommy, there is the chair”, “Mommy is in the chair”, or some other possible interpretation. Bloom’s criticism reflected the emphasis during the 1970s (Leonard 1976, Schlesinger 1974) on the idea that children’s early word combinations were based on the use of some small set of universal conceptual relations such as modifier + modified, locative + locations or subject + verb. In an attempt to align his earlier theory with this Zeitgeist, Braine (1976) suggested that early combinations could best be viewed as “groping patterns” in which the conceptual relations were initially vague, but became solidified over time. Braine viewed patterns of this type as expressing high-level semantic relational features such as recurrence (*another doll*), possession (*my doll*), agency (*doll runs*), or object (*want doll*).

3. Item-based patterns

My own analysis (MacWhinney 1975a) took a somewhat different approach to positional patterns. Rather than arguing that children were selecting combinations from two large classes or expressing a small set of universal conceptual relations, I looked at early combinations as based on an array of what I called “item-based patterns” (IBPs) with each pattern linked tightly to some individual lexical item. This emphasis on generation of syntax from lexical items was in tune with ongoing work at the time on Lexical Functional Grammar (Bresnan 1978, Pinker 1982). Over time, the emphasis on lexical determination of patterns of word combination has increasingly become the default in linguistics, whether it be through the Merge operation (Chomsky 2010) or the feature cancellation of Combinatory Categorial Grammar (Steedman 2000). Because IBPs emphasize individual lexical items as the building blocks of combinations, they avoid the imposition of adult conceptual relations on early child utterances. Instead, the relevant conceptual relations are, at least initially, the relations inherent in individual predicates such as *more*, *want*, or *my*. Rather than viewing the combination *more milk* as expressing a pattern such as recurrence + object, this framework interprets the combination as evidence for the pattern *more* + X, where the italicization of the word *more* indicates that it is a particular lexical item and not a general concept. This analysis stresses the extent to which the IBP first emerges as a highly limited construction based on the single lexical item *more*. These item-based combinations can be viewed as predicate-argument relations. In the IBP for *more milk*, the predicate is *more* and the argument or slot filler is *milk*.

In the case of the IBP for *want* there are two terms that can complete its argument structure. First, there must be a term that serves as a direct object, as in *want cookie*. Often this term is a nominal, but children also produce combinations in which the second term is optionally also a verb, as in *want kiss*. Second, there must be a nominal that serves as the subject, as in *I want cookie*. Because *want* expects these two additional complements, we can call it a two-argument predicate. Other predicates, such as *under* or *my*, take only one argument, and a few such as *give* take three (*John gave Bill a dollar*). The only lexical categories that typically take no additional arguments are nouns, such as *dog* or *justice*, and interjections, such as *gosh* or *wow*. Unlike verbs, adjectives, prepositions and other words that require additional arguments, nouns and interjections can express a full meaning without additional arguments. On the other hand, nouns that are derived from verbs, such as *lack*, *destruction* or *decline* can take prepositional phrases as additional complements (as in *a lack of resources*, *the army's destruction of the city* or *a decline in the dollar*), but basic nouns such as *chair* and *goat* do not even have these expectations for additional complements.

3.1 How children learn IBPs

Children learn item-based patterns by listening to phrases, short sentences, or fragments of longer sentences. For example, if the child's older sister says *this is my dollie*, the child may only store the last two words as *my dollie*. Within this sequence, the child will then recognize the word *dollie* from previous experience and associate that word with the actual doll. This extraction of the "known" segment then leaves the segment *my* as "unknown" or uninterpreted (MacWhinney 1978). At this point, the child can compare the phrase *my dollie* with the single word *dollie*, noticing the differences. The first difference is the presence of *my* before *dollie*. The second difference involves the meaning of possession by the speaker. Because this meaning makes no sense without an attendant argument, it is acquired as a predicate that takes on a meaning when combined with its argument. At this point, the child can establish a new lexical entry for *my* and associate it with the meaning of being possessed by the speaker (the older sister). While acquiring this new form, the child also extracts the item-based pattern *my* + X. This means that, right from the beginning, the construction of this new lexical predicate involves a parallel construction of an IBP. In this case, the older sister may be asserting her control over the doll and wrestling it from the younger sister's possession. Thus, the younger child picks up not only the meaning of *my* and its position with respect to its argument, but also the notion of a relation of possession and control between the two words. The important point here is that IBPs are formed directly when new predicates are learned.

It is more accurate to speak of this item-based pattern as combining *my* + object possessed, rather than just *my* + X. The specific shape of the semantic relation here is shaped by the exact activity involved in the child's possessing this particular doll. Embodied relations of this type can be represented within the general theory of Cognitive Grammar (Langacker 1989) and its more specific implementations in the theory of Embodied Cognition (Feldman 2006). From this perspective, we can see the relations between predicates and their arguments in terms of enacted actions, emotions, perceptions and space/time configurations. For example, when a child says *my dollie*, there is a specific reference to the embodied action of holding the doll. Often we can see this even as the child is talking. When the child says *byebye Daddy*, there is a concomitant waving of the hand and the physical experience of seeing the father leave. When the child sees a toy dog fall from the table and says *puppy fall*, there is a linkage to other experiences with falling either by the child herself or by other objects. In all these relations, children are expressing activities and relations for which they have had direct embodied physical contact and experience.

Initially, the pattern of *my* + object possessed is restricted to the words *my* and *dollie* and the relation of possession that occurs between them. However, if the older sister then says *and this is my horsie*, the child can begin to realize that the open slot for the item-based pattern linked to *my* refers potentially to any manner of toy. Subsequent input will teach the child that any object can fill the slot opened up by the operator *my*. Each IBP goes through this type of generalization which I have called "feature superimposition" (MacWhinney 1975b) and which Worden (2002) calls "feature pruning". By comparing or superimposing forms such as *more milk*, *more toys* and *more cookies*, the child can generalize the semantic features of the argument slot. This comparison prunes out features such as [+ solid] or [+ edible] and leaves features such as [+ object] or [+ force].

Parents can promote the child's learning of IBPs by providing appropriate input structures. As soon as the child begins to understand the frame *where's your* +X, parents can ask *where's your nose*, *where's your tummy*, and so on. Then, they can build on this structure by saying *show me your nose, show me your tummy,* and so on. From teaching sequences such as these, the child can pick up the IBP *your* + X. Sokolov (1993) observed that parents's use of these frames increases at the time when children begin to show understanding of the relevant structures. Second language researchers refer to these repetition structures as "build ups", whereas first language researchers refer to them as "variation sets", because they emphasize the variation that arises in the argument slot of a given IBP. Küntay and Slobin (1996) report that roughly 20% of the input to children involves such variation sets, and Waterfall et al. (2010) have shown that the presence of these sets in the input can improve the learning of computational models of language acquisition.

3.2 The structure of IBPs

This view of the learning of IBPs motivates several assumptions regarding how IBPs are structured and function. Specifically, each IBP specifies:

1. the lexical identity of its predicate, which can be either a free or bound morpheme,
2. the possible lexical features of one or more arguments,
3. the position of the predicate vis a vis its arguments, and
4. the conceptual/grammatical relation that holds between the predicate and each argument.

These four components of the IBP are shaped directly during the initial process of learning of predicates. In this regard, we can also analyse the learning of affixes in terms of IBP learning. For example, the learning of the English plural suffix -s can be described through the same learning scenario we used to describe the learning of the IBP for the quantifier *more*. Consider a child who knows the word *dog* and is now trying to understand the word *dogs*. Following MacWhinney (1978), we can assume that the comparison of the known word *dog* with the new form *dogs*, leads to the masking of the shared phonological segments and the isolation of the -s as the "unknown" segment. Similarly, comparison of the referent of *dog* with the current referent of *dogs* leads to the abstraction of plurality as the "unidentified" concept. The linking of the unknown form to the unidentified concept produces a new lexical form for the plural. This new predicate then links the nominal argument *dog* to the pre-predicate slot and establishes a relation of quantification between the suffix and the noun. Because affix-based patterns are so frequent and consistent, children find them very easy to learn. We know that in English (Braine 1963), Garo (Burling 1959), Hungarian (MacWhinney 1976), Japanese (Clancy 1985) and Turkish (Aksu-Koc and Slobin 1985) the ordering of affixes in words is almost always correct, even at the youngest ages. In general, the learning of affixes is parallel to the learning of other lexical predicates. Of course, there are also differences between the two scenarios in terms of the triggering of phonological processes (MacWhinney 1978), but that is a story for another time.

3.3 Clustering

There are three other aspects of IBPs that arise from different processing sources. The first is the property of clustering, which produces the capacity for recursion. Clustering allows a combination of words to occupy an argument slot. For example, in the sentence *I want more milk*, the combination *more milk* functions as a cluster that can fill the object argument slot for the verb *want*. Clustering allows the child to gradually build up longer sentences and a more complex grammar. Examples of potentially infinite cluster types include structures such as (*John's* (*friend's* (*sister's car*))) or *I know* (*that John said* (*that Mary hoped* (*that Jill would go*))). Chomsky has argued (Chomsky 2010, Hauser et al. 2002) that this type of recursive structuring represents a unique adaptation in human evolution determined by a single recent mutation. MacWhinney (2009), on the other hand, argues that recursion is grounded on a wide set of mnemonic and conceptual abilities in higher mammals that achieved more dynamic functioning once humans had developed systematic methods for encoding lexical items (Donald 1991). For our present purposes, what is important is the way in which the child can appreciate the fact that the combination *more milk* functions in a way that is equivalent to the single item *milk*. Whether this is a recent, unique development or an older basic cognitive function is irrelevant for our current purposes.

3.4 Non-locality

A second important property of IBPs is that they can sometimes specify non-local slot fillers. For example, in the sentence *what did you eat*? the argument slot of *eat* is filled by a non-local element in accord with the Active Filler strategy (Frazier and Flores d'Arcais 1989). The fact that children take years learning to control these patterns (Brown 1968, Kuczaj and Brannick 1979) shows that local attachment is the default assumption. However, the system is capable of eventually picking up all of these non-local positional specifications. Apart from active fillers, IBPs can also encode interrupted attachments, such as the sequence of *can* + NP + V operative in phrases such as *can you sing*? Learning of these discontinuous elements begins in contexts which have only one word or short phrase intervening between the elements, as in *can he go*, *je ne vais pas*, or *er hat kein Frühstück genommen*. Once this basic pattern is established, more complex forms can be created through clustering and adjunct attachments. As Chomsky (2007) notes, because non-local patterns go against the basic principle of economy of local attachment, it is likely that they serve other important pragmatic functions, such as the stressing of the new information in a *wh*-question.

3.5 Agreement

A third aspect of IBPs involves the possibility of additional structural context. This additional structural content is triggered primarily through agreement and complementation. In these structures, IBPs require feature agreement not just between the IBP predicate and its arguments, but also between the IBP predicate and the features of other predicates attached to the arguments. One common form of agreement is between the verb and its arguments; another is between the noun and its modifiers. In a phrase such as *he goes*, the verb affix marking the third person singular agrees in person and number not with its head, but with an argument of its head, the word *he*. Often agreement involves two grammatical morphemes that agree across a relation between their respective bases. In a Spanish phrase such as *mis caballos lindos* ('my pretty horses'), the plural suffix on *mis* ('my') and *lindos* ('pretty') agrees not with the base *caballo* ('horse'), but with the suffix *-s* on the head noun *caballos* to which these words bases attach as modifiers. The German phrase *die Mädchen kochen* ('the girls are cooking') shows a similar structure in which the plurality of the definite article *die* agrees with the plurality of the suffix *-en* on the verb *kochen* ('cook'). These configurations must be marked as entries in the IBPs for each of these grammatical morphemes.

IBPs must also occasionally agree with the features of arguments in subordinate clauses. For example, in the Spanish sentence *supongo que venga* ('I imagine he will go'), the word *venga* in the complement clause is placed into the subjunctive because of the selectional restriction of the main verb *suponer* for the irrealis mood. Another classic case of agreement in the child language learning literature involves children's learning of complement structures in sentence pairs such as *John is eager to please* and *John is easy to please*. In the former, the IBP for *eager* specifies that the perspective/subject (MacWhinney 2008c) of the complement clause is also the subject of the main clause. In the latter, the IBP for *easy* specifies that the perspective/subject of the complement clause is some generic participant that pleases John. Children find it difficult to learn these patterns (Chomsky 1969), not only because of the more complicated IBP structures, but also because of the additional perspectival adjustments they require. This ability of IBPs to trace information across relational arcs into subordinate clauses conforms with the notion of degree-zero learnability proposed by Lightfoot (1989). Lightfoot argued that grammatical relations could be learned primarily from main clauses with only a little bit of "peeking" into the verbs of subordinate clauses. As Lightfoot noted, these restrictions on the effective environment for grammatical relations overcome the various complexities imagined in earlier work on learnability of transformational grammar and the so-called "logical problem of language acquisition" (Wexler and Culicover 1980).

4. Processing IBPs

To understand how children build up these complex syntactic structures in both production and comprehension, we need to consider how a syntactic processor can combine words using item-based patterns (along with the feature-based patterns to be discussed later), operating in real time. Most current accounts of real-time syntactic processors use the basic logic found in the Competition Model of MacWhinney (1987). That model specifies a series of steps for the competition between constructions during comprehension:

1. Sounds are processed as they are heard in speech.
2. Closely related words compete for selection based on the support they receive from input sounds.
3. Each selected word activates its own item-based patterns along with related feature-based patterns.
4. Item-based patterns initiate searches for specified slot fillers.
5. Slots may be filled either by single words or by whole phrases. In the latter case, the attachment is made to the head of the phrase.

6. To fill a slot, a word or phrase must receive support from cues for word order, prosody, affixes, or lexical class.
7. If several words compete for a slot, the one with the most cue support wins.

The details of the operation of this parser are controlled by the competitions between specific lexical items and the cues that support alternative assignments. Consider the case of prepositional phrase attachment. Prepositions such as *on* take two arguments: The first argument is the object of the preposition, the second argument is the head of the prepositional phrase (i.e., the word or phrase to which the prepositional phrase attaches). We can refer to argument #1 (arg1) as the local head or endohead and argument #2 (arg2) as the external head or exohead. Consider the sentence *the man positioned the coat on the rack*. Here, the endohead of *on* is *rack* and its exohead (the head of the whole prepositional phrase) could be either *positioned* or *the coat*. These two alternative attachment sites for the prepositional phrase are in competition with each other.

Competition also governs the interpretation of verbs as either transitive or intransitive. Verbs like *jog* that have both transitive and intransitive readings can be represented by two competing lexical entries. When we hear the phrase, *since John always jogs a mile*, we activate the transitive reading. However, if the full sentence then continues as *since John always jogs, a mile seems like a short distance*, then the intransitive reading takes over from the transitive one. For detailed examples of the step-by-step operations of this type of processor consult MacWhinney (1987) or O'Grady (2005).

Sentence production involves the inversion of many of the operations that occur during comprehension. The basic steps are:

1. The speaker formulates an embodied mental model of an activity (McNeill 1979), focusing on the core verbal predicate (MacWhinney 2008c) and its associated nominal starting perspective.
2. Associated with the core predicate can be interactional markers that are often preposed or postposed to the core predicate.
3. Each predicate activates slots for its arguments in accord with IBPs.
4. Arguments may also trigger the activation of further modifiers and clauses and verbs may trigger the activation of adjuncts in accord with IBP structures.
5. As slots become activated, lexical items are activated to fill them.
6. Production begins with preposed interactional forms and topics, sometimes followed by pauses.

7. When the items linked into a phrasal group have all received lexical activation, the articulator can begin producing that phrase, while other processes continue on later phrases.
8. If some items are not ready in time, there can be pauses, retracings, or other disfluencies.

5. Generative power of IBPs

A central goal in child language research is the formulation of a model of grammatical learning that can simulate or "generate" the utterances produced by the child without also generating forms that are clearly improbable or divergent. Of course, the sentences that are actually recorded in a given set of transcripts may be an incomplete representation of all the forms that the child can produce. However, if the sampling is dense enough (Tomasello and Stahl 2004), they can be viewed as a reasonable approximation to what the child is actually capable of producing.

Applying the concept of IBPs (MacWhinney 1975a), I examined the word order of 11,077 utterances produced by two Hungarian children, Zoli and Moni, between the ages of 17 and 29 months. I found that, across the various samples from the two children, between 85% and 96% of the utterances in each sample could be generated by a set of 40 item-based patterns. In the terms of computational linguistics, this is to say that the use of IBPs achieved a "recall" level of between .85 and .96. This analysis did not consider the "precision" or possible overgeneration of IBPs, because the semantic features on the argument slots were configured to make implausible overgeneration impossible. As we will discuss later, this conservative nature of IBPs is a major strength.

Some examples of these patterns in English translation are: X + *too*, *no* + X, *where* + X, *dirty* + X and *see* + X. The IBP model was able to achieve a remarkably close match to the child's output, because it postulates an extremely concrete set of abilities that are directly evidenced in the child's output. Because of this, it does not suffer from the overgeneration problems faced by Pivot Grammar or the problem of finding a set of universal relational forms that can be applied to early combinations in all languages. The details regarding the ages and lengths of the recordings are as follows:

Table 1. Survey of data from two Hungarian children

Period	Age	Hours	Utterances	Mean Length
Zoli I	1;5, 2-5	4	51	1.10
Zoli II	1;6, 29-30	6	228	1.58
Zoli III	1;8, 6-8	8	2675	1.60
Zoli IV	1;10, 0-6	7	1911	1.87
Zoli V	2;0,0-5	6	835	2.58
Zoli VI	2;2,0-3	7	1826	2.50
Moni I	1;11,18-27	8	1478	1.53
Moni II	2;2,0-7	8	576	1.28
Moni III	2;4,16-17	5	797	1.15
Moni IV	2;5,20-23	8	700	1.03

In order to establish evidence for non-chance use of an IBP, we can use exact probabilities from the table of binomial probability distribution. For example, five identical occurrences of the same order of two equally possible outcomes (either XY or YX) reflects existence of a non-chance pattern at the $p < .032$ level of significance. Similarly, seven orders of one type out of nine trials occurs at the p < .02 level. Given a criterion level of p <.05, the 40 statistically significant Hungarian predicates with their English translations were: *jó* (‘okay’), *is* (‘too’), *csak* (‘just’), *kell* (‘is-needed’), *van* (‘is’), *nem* (‘no’), *ne* (‘don’t’), *addide* (‘gimme’), *hol* (‘where’), *mi* (‘what’), *mit* (‘what-accusative’), *hova* (‘where’), *ott* (‘there’), *itt* (‘here’), *az* (‘that’), *azt* (‘that-accusative’) *ez* (‘this’), *ezt* (‘this-accusative’), *tessék* (‘please’), *én* (‘is’), *gyere* (‘come-imperative’), *látod* (‘see-you’), *most* (‘now’), *de* (‘but’), *csunya* (‘ugly’), *piszkos* (‘dirty’), *enyém* (‘mine’), *kicsi* (‘small’), *másik* (‘another’), *tiszta* (‘clean’), *szép* (‘pretty’), *nagyon* (‘very’), *szervusz* (‘hello’), *halló* (‘hello’), *nézz* (‘look’), *azért* (‘because’), *asztán* (‘then’), *akkor* (‘then’), *megyünk* (‘go-we’), *és* (‘and’) and *mindjárt* (‘soon’).

For the majority of these IBPs, the predicate appeared before the argument. The adjectives and determiners were placed before the following noun with unstressed intonation followed by stress on the noun. The interrogatives were placed before their arguments, but received stress. There were five IBPs in which the predicate followed the argument: *jó* (‘okay’), *is* (‘too’), *csak* (‘just’), *kell* (‘is-needed’) and *van* (‘is’). The positional and intonational forms of the IBPs involved here were more limited than in the standard language, probably reflecting limited variability for these patterns in the input, at least in the short sentences understood by the child.

The ability of IBPs to account for or generate large proportions of early child language corpora has also been noted in more recent analyses (Dąbrowska and Lieven 2005, Lieven et al. 2003, Lieven, Pine, and Baldwin 1997). Lieven and colleagues noted that the construction of an accurate generative corpus based on the child’s utterances works best if the sampling procedure is fairly dense. In the 1997 study, IBPs were able to account for 60% (recall) of the corpus. However, in the 2003 study with denser data, generativity (recall) rose to nearly 90%.

The specific conception of IBPs used in these studies is somewhat different from the one presented here. Lieven and colleagues consider long phrases such as *let’s move it* to be predicates in IBPs. In related work (Bannard, Lieven, and Tomasello 2009), they also allow for discontinuous IBPs, even at early stages in development. Their method treats all frequently occurring word strings as items with slots for multiple possible filler arguments. This emphasis may arise from the fact that Lieven and colleagues tend to focus more on the specific surface strings located in corpora and less on the shape of the grammatical relations used to build up larger units through clustering or combination. It is likely that both lexical and phrasal IBPs contribute to syntactic development, albeit in somewhat different ways (Nelson, 1973).

5.1 Errors as evidence for IBPs

Early child syntax is replete with examples of errors produced by the simple application of item-based patterns (Brown, Cazden, and Bellugi 1969, Klima and Bellugi 1966, Menyuk 1969). Examples include *where Mama boot*, *no Rusty hat* and *that no fish school*. These combinations arise from the application of item-based patterns such as *where* + object located, or *no* + object denied. Eventually, children learn to use *where’s* or *where is*, rather than *where* for interrogatives, producing correct combinations, such as *where’s the wheel*? In these patterns, the open slot can hold single nouns, noun phrases, or simple sentences. When the initial combination of *wh*-word and auxiliary is learned, it is not closely linked to the rest of the discontinuous pattern. This results in double-auxiliary errors (Rowland and Pine 2000) such as *why is the boy can’t come*? Over time children will learn to add these additional components and to unify across item-based patterns. However, while they are occurring, these errors provide direct evidence of the simple combinatorial way in which children are building up utterances from combinations of IBPs.

Some children form an overgeneralized *no* + X negation pattern in which X is not restricted to an object. Errors illustrating this incorrect over-extension include *no do this*, *no wipe finger*, *no sit there*, *no play that*, *he no bite you* and *I no taste them*. Parallel interrogative combination errors include *where go*, *what happen*, *where put him on a chair*, *what happen me* and *why need them more*. Interrogative errors with misplaced auxiliaries of the shape *what they are doing* and *where he's going* are extremely common. There are also errors, such as *where the wheel do go* and *what you did eat*, in which the auxiliary is misplaced after the subject. These errors are further evidence for patterns such as *where* + S. Later on, children replace the IBP *where* + S with the more general pattern of *where* + tense. However, they fail to restrict the *where* + tense pattern to exclude main verbs. Overgeneralization errors attesting to the productivity of this later pattern include *where goes the wheel*, *where could be the shopping place*, *where's going to be the school*. After the first few months of word combination, there are no reports of errors that go against the basic IBP interrogative patterns. For example, there are no reports of errors such as *he can't do it why* (Labov and Labov 1978).

The fact that seemingly general grammatical patterns are often acquired word by word provides further evidence for the operation of IBPs. For example, Kuczaj and Brannick (1979) showed that children are quicker to show placement of the tensed auxiliary after the interrogatives *what* and *where* than after *how long* or *when*. Thus, children will produce *what is he doing*? at the same time they produce *when he coming*? Similarly, Bowerman (1978) noted that, at 17 months, her daughter Eva used the patterns *want* + X and *more* + X productively. However, these patterns did not generalize to other words like *open*, *close*, *bite*, *no more*, or *all gone*. Also, Pine and Lieven (1997) showed that, within the determiner class, the indefinite article appeared in very different combinations from the definite article. For example, some children used the frame *that's a* + X, but not the frame *that's the* + X. In general, the definite article appeared more frequently than the indefinite in prepositional contexts such as *in the* + X. In addition, there were many errors with indefinites such as *a my car*, *a pants*, or *a that*, but no similar errors for the definite. In general, it appeared that the early acquisition of these two forms involved separate IBPs.

One could argue that these various errors are produced not through word combination, but through analogy. However, accounts based on analogy often predict many error types that never occur. For example, Kuczaj and Brannick (1979) noted that questions like *gonna he go*? have never been reported, although children say *he's gonna go*, *he will go* and *will he go*? If analogy were operating here, we would expect to find *gonna he go*? on analogy with *will he go*? However, IBPs account for these data correctly. The auxiliary *will* is combined with *he go* using the IBP *will* + action. This pattern does not generalize to *gonna*, because, by definition, the IBP will + action is restricted to the auxiliary *will*. The item *gonna* never appears in initial position without a preceding nominal, so there is no evidence or form in support of an error such as *gonna he go*?

5.2 Conservatism as evidence for IBPs

Because IBPs are linked to individual lexical items, they do not initially generalize to larger groups of words. We can refer to the fact that IBPs have a limited scope as lexical conservatism. IBP learning involves an ongoing process of generalization for the semantic features of the arguments. During these processes of generalization, to minimize the possibility of error, the child has to be conservative in three ways:

1. The child needs to formulate each syntactic combination as an IBP.
2. Each IBP needs to record the exact semantic status of each positive instance of an argument in a particular grammatical configuration (MacWhinney 1987).
3. Attempts to use the item-based construction with new arguments must be closely guided by the semantics of previously encountered positive instances.

If the child has a good memory and applies this method cautiously, overgeneralization will be minimized and there will be no need to recover from overgeneralization.

IBPs support gradual, conservative productivity. We can also demonstrate the productivity of IBPs by teaching children novel words that serve as argument slot fillers. For example, we can show a child a picture of a birdlike creature that we call a wug. The positioning of the nonce word *wug* after the article *the* induces the child to treat the word as a common noun. We then show the child two pictures of the strange creature and ask them, "What are these?" By responding with the answer *wugs*, children show productivity of the IBP based on the plural suffix. Similarly, we can set up a game in which each person owns some toys. This will lead the child to produce the combination *my wug*, thereby showing the productivity of the pattern *my* + object possessed. Also, a German-speaking child can be taught the nonce name *der Gann* (nominative, masculine and singular) for a toy. The experimenter can then pick up the toy and ask the child what he is holding. By the age of 3, children will correctly produce the accusative form *den Gann* (accusative, masculine and singular). Here, the IPB for the article *den* is responding correctly to the masculine gender on the argument. Although it is easy to convince children to accept new slot fillers, it is far more difficult to teach them entire new IBPs for new predicates. As a result, it is difficult to convince children to use novel verbs in a fully productive fashion. Instead, children tend to be conservative and unsure about how to use verbs productively until past age 3 (Tomasello 2000).

Conservatism helps the child avoid overgeneralization. For example, a non-conservative learner might notice that both *big* and *red* pattern together in forms such as *big barn* and *red barn*. This might induce the learner to produce forms such as *I painted the barn big* on analogy with *I painted the barn red*. However, a more conservative learner would stick close to facts about the verb *paint* and the arguments that it permits (Perfors, Tenenbaum, and Wonnacott 2010). If the child has heard a form like *I painted the barn white*, it would make sense to extend the IBP for *paint* to also allow for *red* in one of the argument slots of the IBP. However, to extend from the word *white* to semantically unrelated words like *happy* or *difficult* would be to go far beyond the attested construction and would not be in accord with the principle of superimposition or feature pruning. In fact, when we examine corpora, we find that this type of category-leaping overgeneralization is extremely infrequent.

The roots of lexical conservatism can be traced back to the earliest periods of word learning. As MacWhinney (1975a) demonstrated in his analysis of the productions for Zoli and Moni, early predicates are very limited in scope. Tomasello (1992) provided detailed support for this claim in his diary study of his daughter's early verbs. From these diary observations, he formulated the "verb island" hypothesis which holds that verbs are initially acquired in ways that are limited to particular syntactic frames and particular usage contexts. This verb island hypothesis is entirely in accord with the theory of IBPs. Later, Brooks and Tomasello (1999) showed experimentally that, when 3-year-olds were taught novel verbs, they conservatively avoided treating transitives as intransitives, despite the fact that there is a process in English that permits this generalization.

The acquisition of morphological markings for declension and conjugation provides further evidence for conservatism. For example, young Spanish children may be quick to produce the first person present of *tener* 'have' as *tengo*, but this does not immediately generalize to producing the first person plural *tenemos*, even if the child uses other first person form forms like *comemos* 'we eat' or *jugamos* 'we play'. In other words, the IBP for the first plural suffix *-mos* is initially bound to specific forms that can fill the argument slot. Eventually this slot generalizes, but that generalization can take a few days or even weeks.

5.3 Correlational evidence for IBPs

The theory of IBPs stipulates a nearly complete linkage between lexical learning and syntactic learning during the first year of language development. This linkage arises from the claim that the learning of predicates leads directly to the formulation of the IBPs that combine predicates with their arguments. Evidence for this linkage of syntactic learning with lexical learning is provided in the analysis of the first stages of language learning performed by Bates and Goodman (1999). In their analysis of lexical and syntactic development during the first year of language learning, Bates and Goodman found a correlation of 0.96 between lexical and syntactic developments through the first year. Moreover, the size of the lexicon at age 2;0 was an excellent predictor of the level of syntactic development at age 3;0. Bates and Goodman argued that this evidence pointed to the fundamental inseparability of grammar and the lexicon during this period. This tight relation is exactly what we would expect if the process of lexical acquisition leads directly to advances in syntactic combinations through the establishment of IBPs.

Goldberg (2006: 18) holds that "it's constructions all the way down." According to this view, morphological patterns, lexical items, phrasal patterns and complex sentence structure are all instances of constructions. Given this view, it is easy to imagine that all of grammar could be erected on a lexical basis. It would also be tempting to imagine that this tight correlation between lexicon and grammar would continue throughout development. Such a tight correlation is generally consistent with the theory of IBPs. However, as I have argued elsewhere (MacWhinney 2009), grammar eventually achieves an independent status, although it continues to cooperate and compete intimately with the lexicon throughout adulthood.

5.4 Competition and conservatism

When learning the item-based construction for ‘give’, children encounter sentences such as *Bill gives John the book*. From this, they learn the double-object IBP: giver + ‘give’ + recipient + gift. They also learn the competing periphrastic IBP of giver + ‘give’ + gift + ‘to’ recipient. Moreover, because both of these IBPs are linked to the same verb, they are in direct competition. Therefore, we need to consider the explanatory mechanisms proposed in the Competition Model (MacWhinney 1987) for understanding how choices are made between competing IBPs. In this case, the winner of a given competition is the one that receives further support from cues such as focusing or topicalization. Some verbs, such as *donate* or *recommend*, only allow the periphrastic IBP. In those cases, because there is no competition, conservatism solves the problem directly. The principle of conservatism is equivalent to what others have called the Subset Principle (Fodor and Crain 1987). According to the Subset Principle, more restrictive grammars are always favored over more general grammars. By linking the principle of conservatism with the principle of competition, we derive additional descriptive power not available from simple, uniform application of the Subset Principle alone. In particular, the Subset Principle cannot account for occasional overgeneralizations, whereas conservatism with competition can explain how children are generally conservative, but occasionally willing to overgeneralize.

6. Feature-based patterns

Although IBPs can be used to generate nearly all word combinations, children eventually go beyond IBPs to formulate more general combinatorial rules, based on grammatical relations between major part of speech categories (MacWhinney 1975a, 1982). Consider the learning of the pattern that places the adjective before the noun in English. At first, children pick up a few IBPs such as *nice* + object, *good* + object and *pretty* + object. They acquire these patterns during the learning of new adjectives from the input. For example, children may hear the form *nice kitty*, from this they learn the word *nice*, while also establishing the IBP *nice* +X. At first, the argument slot is limited to the original noun *kitty*. However, after hearing forms such as *nice baby* and *nice party* the argument slot begins to be open to all possible objects. When the child then learns IBPs for *good* + X and *pretty* +X, the process of slot generalization becomes quicker, as the child begins to realize that words like *nice*, *good* and *pretty* that describe characteristics of objects all accept a related object in the following syntactic position. This linking of collections of similar item-based patterns then gives rise to a feature-based pattern or FBP (MacWhinney 1982) that specifies the combination modifier + modified for English. The learning of the modifier class is further promoted by the acquisition of an additional FBP for postcopular positioning of predicate adjectives.

Other early FBPs include possessor + possession (*John's computer*) and locative + location (*behind the tree*). Once children have learned these more general patterns, they apply them immediately to newly learned words. This account of the emergence of FBPs is very different from that found in the proposals from the 1970s by Braine, Schlesinger and others. Those proposals viewed FBPs as available to the child during the first steps of language learning. MacWhinney (1975a) showed that this first step involved the learning of IBPs, not FBPs, thereby addressing Bloom's (1971) critique of Pivot Grammar and other approaches that assumed that children's first grammatical patterns involved large word classes.

In languages like English or German, where the adjectives and other modifiers reliably precede the noun and occur in consistent positions after the copula, the acquisition of the FBP for modifier + modified is probably the first FBP to reach full productivity. In Hungarian, the acquisition of this pattern is a bit slower, because of the fact that copula omission makes it more difficult to unify the prenominal modifier frame with the frame of predicative adjectives. Another early FBP in English is for the subject or perspective (MacWhinney 2008c) of the verb. In English, the positioning of a noun before the verb is a remarkably reliable cue to the role of perspective or first argument. Again, the reliability of this pattern promotes its early acquisition as a FBP.

6.1 Evidence for FBPs

There are three forms of evidence for the functioning of FBPs. As in the case of IBPs, we can find evidence for the learning of FBPs from studies of child language corpora. For example, by the age of 3, children make use of the adjectives they know in both prenominal and predicate adjective position. Second, we can demonstrate the productivity of FBPs by studying the treatment of newly learned or nonce words. For example, with a five-year-old, an experimenter can introduce a new verb like *griff* in the frame *Tim griffed the ball to Frank* and the child will productively generalize to *Tim griffed Frank the ball*. Third, we can find evidence for the productivity of FBPs by tracing the production of syntactic overgeneralization errors, such as *he said me the story* or *pour the tub with water*. For example, tracing the production of verb constructions such as the double object, the passive and the causative, Bowerman (1988) demonstrated productivity of these constructions during the period from 4 to 8, but not before. MacWhinney (2004) shows how these errors demonstrate a competition between IBPs and FBPs for the relevant verb forms.

6.2 Global patterns

The predicates and arguments being related by FBPs are specified both in terms of their grammatical relation and in terms of their part of speech features. However, there are also patterns that apply still more generally across part of speech categories. We can refer to these more general FBPs as global patterns. One such pattern is the relation between a topic and a comment. In this structure, the topic can be a noun, adjective, or prepositional phrase. The comment can be a verb, noun, or adjective. These structures occur frequently in languages like Hungarian or Chinese. In these languages topics are placed in initial position, followed by an intonational break and then the comment. When learning this first, children pick up this pattern in the context of a few IBPs. For example, they might hear a Hungarian sentence of the shape *the glass # empty* with the # sign indicating an intonational break between the topic and the comment. They first encode this as a pattern linked to *glass*. However, after hearing a few other similar patterns for other nouns, they then extract a FBP, just as they do for the modifier + object described pattern for adjectives. Studies such as MacWhinney (1975a) and Lee (1999) have demonstrated that children use these patterns productively by age 2. However, learning of this type of pattern can also continue further, as children find that not only nouns, but also other structures, can serve as topics.

Global patterns are also involved in the configuration of basic word orders such as SVO for English. Akhtar and Tomasello (1997) provide an interesting type of evidence regarding the learning of such overall sentence templates. They found that two-year-olds and three-year-olds were willing to repeat sentences using novel verbs in the non-standard SOV and VSO orders, along with standard English SVO. However, by age 4, children assimilated the non-standard orders to standard SVO order. This indicates a gradual imposition of the overall global pattern for SVO order by age 4.

6.3 FBPs and the syntactic component

Children's learning of FBPs and global patterns leads to the construction of a syntactic system that is no longer fully dependent on lexical control through IBPs. In neurological terms, this control is centered in the interior frontal gyrus (IRG), whereas lexical processing arises largely from the medial temporal gyrus (MTG) (Friederici 2009). MacWhinney (2009) argues that the emergence of frontal support for grammar is important, because it allows for smoother communication with frontal mechanisms for perspective taking and mental model construction. Although IBPs are capable in principle of controlling sentence production and comprehension, development of a separate inferior frontal gyrus (IFG) system for linear control of syntactic patterns can promote advances in linguistic fluency and complexity. FBPs allow for the uniform encoding of proceduralized linear structures that operate in computational terms like finite-state automata (Pulvermüller 2003). However, these patterns must still rely on accurate reciprocal communication with the posterior lexical component in order to avoid problems with disfluencies and other speech errors (Stemberger 1985).

In order to achieve effective communication with the lexicon, the emerging syntactic component must rely on white matter tracts that connect the relevant anterior and posterior areas. Recent work using DTI (diffusion tensor imaging) has shown that these tracts maintain careful patterns of connectivity between these two areas, as well as between both superior temporal gyrus (STG) and medial temporal gyrus (MTG) posterior areas (Bookheimer 2002). Communications across these white matter tracts require that a signal from anterior regions for the activation of a given element in a given part of speech match up with the activation of the relevant word in posterior lexical regions. The DevLex model (Li, Zhao, and MacWhinney 2007) shows how this communication can be facilitated through the emergence of local regions in lexical space that encode topologically for given parts of speech. In this way, the communication between syntax and lexicon can be maximally accurate. A goal for future neural network modeling is to show how the DevLex model integrates with IBPs and FBPs over time.

7. Item-based patterns in second language learning

In section 4 we examined the linkage between IBP theory and the processing account of the Unified Competition Model or UCM (MacWhinney 2012). Within that context, IBP theory serves as the mechanism driving syntactic learning and linking syntax to lexicon. According to the UCM, when we turn our attention to second language (L2) learning, we have to take into account the impact of the additional factors of transfer, parasitism, entrenchment and resonance. Although a complete account of these processes in L2 learning is beyond the scope of this paper, we can consider the general way in which IBPs could play a role in this process. In this area, the UCM claims that "everything that can transfer will". However, transfer is only predicted if there is a good alignment between the structure of the source and the target. If this alignment is close, we could expect to see some transfer of first language (L1) patterns for IBPs. Pienemann et al. (2005) noted that, despite the fact that both Swedish and German have a pattern that places verbs after initial adverbial phrases (as in *heute kommt er* ('he is coming today')), Swedish learners of German are reticent to transfer this pattern en masse. Instead, they seem to begin back at the level of the IBP, picking up this adverb-initial pattern one adverbial type at a time. This suggests that second language learners employ the same type of IBP conservatism that we observed in children. In part this may be a necessary property of a language learning system. Without first grounding learning on a database of confirmed IBPs, it may be impossible or even risky to establish a full set of FBPs. However, as learning of IBPs progresses, the database becomes rich enough to support a nativelike FBP system, contrary to the claims of Clahsen and Felser (2006) regarding the inability of second language learners to acquire nativelike syntax. Of course, second language learning takes time, but eventually the learner gets closer and closer to the native model. The smooth, gradual nature of the development of nativelike cue strength has been documented in Competition Model studies of learners of Dutch and English (Kilborn and Cooreman 1987, McDonald 1986).

In another regard, L2 IBP learning is quite different from L1 IBP learning. In the case of L1, the learning of the IBP is closely related to the learning of the predicate itself as a new lexical item. In the case of L2, there is also conceptual transfer from the L1 translation equivalent to the new L2 predicate. For example, when learning the German word *viele* 'many', it is easy for an English-speaking learner to simply transfer the core IBP for 'many', along with its relational expectations and lexical semantics. For

German, this could then also lead to a transfer of the agreement of *viele* with plurality on its argument noun without producing any errors. However, when learning Hungarian, a transfer of the English IBP to the Hungarian word *sok* 'many' would lead to a conflict, because Hungarian disallows marking of the plural in this context. In fact, there is good evidence that many forms of interlanguage errors can be explained through IBP transfer in this way.

8. Comparisons with other models

The theory of item-based patterns shares assumptions with a number of linguistic approaches. In the American tradition, the emphasis on lexical determination of syntax was a central feature of Bresnan's (1978) work on lexical-functional grammar (LFG). This emphasis was continued in HPSG (Pollard and Sag 1994) and Construction Grammar (Goldberg 1995, 2006, Langacker 1986). Within the European tradition, there is a long history of development of Categorial Grammar and Dependency Grammar (Hudson 1984, Mel'cuk and Polguère 1987, Tesniére 1959). The models in this tradition that are closest to work on IBPs and the Competition Model are the left-associative grammar (LAG) of Hausser (1992, 1999) and the Unification Space model of Vosse and Kempen (2000). In work with the GRASP model (Sagae et al. 2010, Sagae et al. 2007, Sagae, Lavie, and MacWhinney 2005, Sagae, MacWhinney, and Lavie 2004a, 2004b), we have elaborated a specific Dependency Grammar model that has been used to tag the grammatical relations in the English, Spanish, Mandarin and Hebrew segments of the CHILDES database (MacWhinney 2000). However, GRASP makes relatively little use of item-based patterns, emphasizing instead the role of larger word classes for determining grammatical relations. This is done because GRASP is designed to achieve accurate tagging and is not designed to account for the course of child language acquisition.

Recently, there have been several computational models that have addressed the core developmental issue, often by focusing on the statistical properties of the input to the child. However, it appears that none of these approaches has yet implemented the core ideas of the IBP approach. By way of summary, we can highlight these four claims from IBP theory:

1. Children can acquire new words from simple sentences in which words appear either alone or in small phrasal combinations.

2. The learning of a new predicate from analysis of a combination leads directly to the formulation of a new IBP.
3. IBPs arguments are generalized by pruning.
4. When possible, IBPs are combined into FBPs.

Current models of syntactic learning tend to skip over the steps emphasized in the second and third claim, focusing instead on the direct establishment of word classes without attention to an intermediate stage of IBP functioning. These current approaches can be grouped into six types, based on the assumptions they make about the input to the learner. In all but one of these models, the default assumption is that an induction procedure should take as its input the utterances produced by parents and yield as its output the grammar used by the child.

1. Learning from full scenes and utterances. Working within the framework of Machine Learning theory, Mooney and colleagues (Ge and Mooney 2005, Mooney 2008, Zelle and Mooney 1993) have emphasized ways in which descriptions of whole scenes can be aligned with full sentences to extract words and grammar. This whole-scene learning approach has also been elaborated by Siskind (2000) and Roy (2002). These approaches hold great promise for Machine Learning. However, they make the unrealistic assumption that children achieve complete storage of all input and subsequent batch processing of a full database to extract regularities. This framework ignores both the child's memory limitations and the incremental nature of the learning process. Recent unpublished work by Deb Roy on his son's learning of the word for *water* (Roy, personal communication) is an exception to that pattern. However, it is not yet clear how this work on lexical learning will hook up with theories of syntactic learning.
2. Learning from tagged corpora. Two recent language learning models have taken as their starting point the grammatical tags introduced into the CHILDES corpora by the GRASP tagging system (Sagae et al. 2010). The model of Borensztajn, Zuidema and Bod (2009) relies on the Unsupervised Data-Oriented Parsing (U-DOP) framework of Bod (2009), coupled with a Hierarchical Prediction Network (HPN). Another system of this type is being developed by Mark Steedman and colleagues (Kwiatkowski et al. 2010), using the formalism of Combining Categorial Grammar (Steedman 2000) and processes of Bayesian learning (Tenenbaum, Griffiths, and Kemp 2006). Both of these approaches are able to show high levels of learning. Their successes

show that claims that the input to the child is too impoverished to support language learning cannot be right (MacWhinney 2004). However, these models achieve their success by relying on pre-existing codes for grammatical relations in the CHILDES database. One could argue that these relations approximate relations that would be observable in the situational context. However, in reality, these relations are extracted by relying on a previous stage of part of speech tagging (MacWhinney 2008b) that involves information not yet fully available to the language learning child. IBP theory approaches this issue by emphasizing the role of incremental organization of part of speech information within the self-organizing DevLex (Li, Zhao, and MacWhinney 2007) neural network. That approach seems to provide a closer model of the ways in which lexical learning supports syntactic organization.

3. Learning from constituent extraction. Harris (1951) emphasized the idea that linguistic patterns could be extracted through constituent structure analysis. The ConText model of Waterfall, Sandbank, Onnis and Edelman (2010) uses distributional clustering from CHILDES data to produce equivalence classes. These classes are configured within a probabilistic context free grammar (PCFG) to produce sentences designed to match the child's output. The results of this system were not as accurate as the systems that learn from tagged corpora, although there was clear evidence of consistent learning. Going beyond these basic results, Waterfall et al. show that additional, more realistic, learning can be achieved by providing input from variation sets of the type discussed in section 3.1. If work on ConText could also integrate ideas about item-based learning, the account would match up still better to the steps followed by children.
4. Modeling the child. Bannard, Lieven and Tomasello (2009) take an approach to learning that is very different from those we have discussed so far. Rather than inducing a child grammar from the adult input (possibly supplemented by situational information), they compose a grammar based on the child's own utterances. Like ConText and the IBP approach, they assume a probabilistic context-free grammar (PCFG). The learning method they use relies on a process of alignment of full utterances to extract areas of overlap and areas of mismatch. In the terms of the current analysis, the areas of overlap are treated as predicates for new IBPs and the areas of mismatch are treated as the open argument slots for these items. Basically, they are trying to reason backwards from the child's corpus to the underlying

IBPs. However, unlike the IBP theory developed here, they allow for many forms of discontinuous constituents and multi-lexical forms. Of course, such forms may eventually play a role in processing, but including them by default from the earliest time seems to underestimate the actual role of individual lexical items as the controlling core of IBPs. Also, like most other current models, this approach involves overly strong assumptions regarding batch processing and the absence of memory limitations. However, further work in this framework may be able to relax some of these assumptions.

5. Dealing with specific items. Among current computational approaches, the one that aligns most closely with the IBP theory developed here is the account of the learning of the English dative developed by Perfors, Tenenbaum and Wonnacott (2010). This account used Hierarchical Bayesian Modeling (HBM) to extract successive levels of generality. This process seems closely related to the shift from IBPs to FBPs and eventually global patterns developed in the current analysis. Although this HBM account targets only a small component of the grammar, it focuses clearly on the ways in which the statistics of the input can be used to cluster dative verbs into types, much as suggested by Braine (1971) in his hierarchical sieve model.
6. Setting parameters. The theory of Principles and Parameters of P&P (Chomsky and Lasnik 1993) holds that the learning of syntax involves nothing more than the setting of the correct values on a few binary structural parameters. For example, children must set the correct value on the Null Subject parameter in order to decide whether or not it is permissible to drop subjects. Italian-speaking children must learn that this is possible, whereas English-speaking children must learn that it is not (Hyams 1986, Hyams and Wexler 1993). Although P&P foregrounds some interesting patterns in language typology, it fails to provide detailed accounts of the specific distribution of utterances during language development. To bridge this gap, theorists often invoke the performance/competence distinction, holding that children begin language learning with full competence (Poeppel and Wexler 1993) and that deviations from adult grammar are due only to performance constraints. However, Braine (1974) showed that performance limitations fail to provide proper characterizations of the utterances actually produced by children.

9. Conclusion

Although models still differ in terms of specific assumptions, work on early syntactic development is making increasingly productive use of the theory of item-based patterns. The availability of large, dense corpora (Lieven et al. 2003), the automatic tagging of corpora for part of speech and grammatical relations (Sagae et al. 2010) and the deployment of well-formalized computational methods are leading to rapid advances in our understanding of the mechanisms of syntactic learning. The major challenge we now face is understanding how best to integrate a data-driven approach that learns on the basis of the parental input with a more descriptive approach that models the child's productions directly (Bannard, Lieven, and Tomasello 2009, MacWhinney 1975a).

References

Ahktar, Nameera, and Michael Tomasello
1997 Young children's productivity with word order and verb morphology. *Developmental Psychology* 33: 952–965.
Aksu-Koc, Ayhan, and Dan Isaac Slobin
1985 The acquisition of Turkish. In *The Crosslinguistic Study of Language Acquisition*, Dan Isaac Slobin (ed.), 839–880. Hillsdale, NJ: Lawrence Erlbaum Associates.
Aslin, Richard N., Jenny R. Saffran, and Elissa L. Newport
1999 Statistical learning in linguistic and nonlinguistic domains. In *The Emergence of Language*, Brian MacWhinney (ed.), 359–380. Mahwah, NJ: Lawrence Erlbaum Associates.
Augustine, Saint (Saint Augustine)
1952 *The Confessions*, vol. 18. Chicago: Encyclopedia Britannica.
Bannard, Colin, Elena Lieven, and Michael Tomasello
2009 Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences of the United States of America* 106: 17284–17289.
Bates, Elizabeth, and Judith Goodman
1999 On the emergence of grammar from the lexicon. In *The Emergence of Language*, Brian MacWhinney (ed.), 29–80. Mahwah, NJ: Lawrence Erlbaum Associates.
Bates, Elizabeth, and Brian MacWhinney
1987 Competition, variation, and language learning. In *Mechanisms of Language Acquisition*, Brian MacWhinney (ed.), 157–194. Hillsdale, NJ: Lawrence Erlbaum Associates.

Blanchard, Daniel, Jeffrey Heinz, and Roberta Golinkoff
2010 Modeling the contribution of phonotactic cues to the problem of word segmentation. *Journal of Child Language* 37: 487–511.
Bloom, Lois
1971 Why not pivot grammar? *Journal of Speech and Hearing Disorders* 40: 40–50.
Bloom, Lois
1973 *One Word at a Time: The Use of Single Word Utterances*. The Hague: Mouton.
Bod, Rens
2009 Constructions at work or at rest. *Cognitive Linguistics* 20: 129–134.
Bookheimer, Susan
2002 Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience* 25: 151–188.
Borensztajn, Gideon, Willem Zuidema, and Rens Bod
2009 Children's grammars grow more abstract with age: Evidence from an automatic procedure for identifying the productive units of language. *Topics in Cognitive Science* 1: 175–188.
Bowerman, Melissa
1978 Systematizing semantic knowledge: Changes over time in the child's organization of word meaning. *Child Development* 49: 977–987.
Bowerman, Melissa
1988 The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? In *Explaining Language Universals*, John Hawkins (ed.), 73–104. London: Blackwell.
Braine, Martin
1963 On learning the grammatical order of words. *Psychological Review* 70: 323–348.
Braine, Martin
1966 Learning the position of words relative to a marker element. *Journal of Experimental Psychology* 72: 532–540
Braine, Martin
1971 On two types of models of the internalization of grammars. In *The Ontogenesis of Grammar: A Theoretical Symposium*, Dan Isaac Slobin (ed.), 153–186. New York: Academic Press.
Braine, Martin
1974 Length constraints, reduction rules, and holophrastic processes in children's word combinations. *Journal of Verbal Learning and Verbal Behavior* 13: 448–456.

Braine, Martin
1976 Children's first word combinations. *Monographs of the Society for Research in Child Development* 41 (whole no. 1).
Branigan, George
1979 Some reasons why successive single word utterances are not. *Journal of Child Language* 6: 411–421.
Bresnan, Joan
1978 A realistic transformational grammar. In *Linguistic Theory and Psychological Reality*, Morris Halle, Joan Bresnan and George Miller (eds.), 1–59. Cambridge, MA: MIT Press.
Brooks, Patricia J., and Michael Tomasello
1999 How children constrain their argument structure constructions. *Language* 75: 720–738.
Brown, Roger
1968 The development of Wh acquisition in child speech. *Journal of Verbal Learning and Verbal Behavior* 7: 279–290.
Brown, Roger, Courtney Cazden, and Ursula Bellugi
1969 The child's grammar from I to III. In *Minnesota Symposia on Child Development*, vol. II, John P. Hill (ed.), 28–73. Minneapolis: University of Minnesota Press.
Burling, Robbins
1959 Language development of a Garo and English speaking child. *Word* 15: 45–68.
Chomsky, Carol Schatz
1969 *The Acquisition of Syntax in Children From 5 to 10*. Cambridge, MA: MIT Press.
Chomsky, Noam
2007 Approaching UG from below. In *Interfaces + Recursion = Language?*, Uli Sauerland and Hans-Martin Gaertner (eds.), 1–30. New York: Mouton de Gruyter.
Chomsky, Noam
2010 Some simple evo devo theses: How true might they be for language. In *The Evolution of Language: Biolinguistic Perspectives*, Richard Larson, Viviane Déprez and Hiroko Yamakido (eds.), 45–62. Cambridge: Cambridge University Press.
Chomsky, Noam, and Howard Lasnik
1993 The theory of principles and parameters. In *Syntax: An International Handbook of Contemporary Research*, Joachim Jacobs (ed.), 1–32. Berlin: Walter de Gruyter.
Clahsen, Harald, and Claudia Felser
2006 Grammatical processing in language learners. *Applied Psycholinguistics* 27: 3–42.

Clancy, Patricia M.
1985 The acquisition of Japanese. In *The Crosslinguistic Study of Language Acquisition: Vol. 1. The Data*, Dan Isaac Slobin (ed.), 373–524. Hillsdale, NJ: Lawrence Erlbaum Associates.
Dąbrowska, Eva, and Elena Lieven
2005 Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* 16: 437–474.
Donald, Merlin
1991 *Origins of the Modern Mind*. Cambridge, MA: Harvard University Press.
Feldman, Jerome
2006 *From Molecule to Metaphor: A Neural Theory of Language*. Cambridge, MA: MIT Press.
Fodor, Janet, and Stephen Crain,
1987 Simplicity and generality of rules in language acquisition. In *Mechanisms of Language Acquisition*, Brian MacWhinney (ed.), 35–63. Hillsdale, NJ: Lawrence Erlbaum.
Frazier, Lyn, and Giovanni B. Flores d'Arcais
1989 Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language* 28: 331–344.
Friederici, Angela
2009 Brain circuits of syntax: From neurotheoretical considerations to empirical tests. In *Biological Foundations and Origin of Syntax*, Strüngmann Forum Reports, Vol. 3, Derek Bickerton and Eörs Szathmáry (eds.), 239–252. Cambridge, MA: MIT Press.
Ge, Ruifang, and Raymond Mooney
2005 A statistical semantic parser that integrates syntax and semantics. Paper presented at the CoNNL-2005, Ann Arbor.
Gleitman, Lila
1990 The structural sources of verb meanings. *Language Acquisition* 1: 3–55.
Goldberg, Adele
1995 *Constructions*. Chicago: University of Chicago Press.
Goldberg, Adele
2006 *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
Harris, Zellig S.
1951 *Structural Linguistics*. Chicago: University of Chicago Press.
Hauser, Mark, Noam Chomsky, and Tecumseh Fitch
2002 The faculty of language: What is it, who has it, and how did it evolve? *Science* 298: 1569–1579.

Hausser, Roland
1992 Complexity in left-associative grammar. *Theoretical Computer Science* 106: 283–308.
Hausser, Roland.
1999 *Foundations of Computational Linguistics: Man-Machine Communication in Natural Language*. Berlin: Springer.
Hudson, Richard
1984 *Word Grammar*. Oxford: Blackwell.
Hyams, Nina
1986 *Language Acquisition and the Theory of Parameters*. Dordrecht: D. Reidel.
Hyams, Nina, and Kenneth Wexler
1993 On the grammatical basis of null subjects in child language. *Linguistic Inquiry* 24 (3): 421–459.
Keenan, Janice, Brian MacWhinney, and Deborah Mayhew
1977 Pragmatics in memory: A study in natural conversation. *Journal of Verbal Learning and Verbal Behavior* 16: 549–560.
Kilborn, K., and Cooreman, A.
1987 Sentence interpretation strategies in adult Dutch-English bilinguals. *Applied Psycholinguistics* 8: 415–431.
Klima, Edward S., and Ursula Bellugi
1966 Syntactic regularities in the speech of children. In *Psycholinguistics Papers*, John Lyons and Roger J. Wales (eds.), 183–208. Edinburgh: Edinburgh University Press.
Kuczaj, Stan, and Nancy Brannick
1979 Children’s use of the Wh question modal auxiliary placement rule. *Journal of Experimental Child Psychology* 28: 43–67.
Küntay, Aylin, and Dan Isaac Slobin
1996 Listening to a Turkish mother: Some puzzles for acquisition. In *Social Interaction, Social Context, and Language: Essays in Honor of Susan Ervin-Tripp*, Dan Isaac Slobin, Julie Gerhardt, Amy Kyratzi and Jiansheng Guo (eds.), 265–286. Hillsdale NJ: Erlbaum.
Kwiatkowski, Tom, Luke Zettlemoyer, Sharon Goldwater, and Mark Steedman
2010 Inducing probabilistic CCG grammars from logical form with higher-order unification. Paper presented at the Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, MIT.
Labov, Teresa, and William Labov
1978 The phonetics of cat and mama. *Language* 54: 316–852.
Langacker, Ronald W.
1986 An introduction to cognitive grammar. *Cognitive Science* 10: 1–40.

Langacker, Ronald
1989 *Foundations of Cognitive Grammar. Vol. 2: Applications*. Stanford: Stanford University Press.
Lee, Thomas Hun-tak
1999 Finiteness and null arguments in child Cantonese. *Tsinghua Journal of Chinese Studies* 33: 1–16.
Leonard, Leonard
1976 *Meaning in Child Language*. New York: Grune and Stratton.
Li, Ping, Xiaowei Zhao, and Brian MacWhinney
2007 Dynamic self-organization and early lexical development in children. *Cognitive Science* 31: 581–612.
Lieven, Elena, Heike Behrens, Jennifer Speares, and Michael Tomasello
2003 Early syntactic creativity: A usage-based approach. *Journal of Child Language* 30: 333–370.
Lieven, Elena, Julian Pine, and Gillian Baldwin
1997 Lexically based learning and the development of grammar in early multi-word speech. *Journal of Child Language* 24: 187–219.
Lightfoot, D.
1989 The child's trigger experience: Degree-0 learnability. *Behavioral and Brain Sciences* 12: 321–275.
MacWhinney, Brian
1975a Pragmatic patterns in child syntax. *Stanford Papers and Reports on Child Language Development* 10: 153–165.
MacWhinney, Brian
1975b Rules, rote, and analogy in morphological formations by Hungarian children. *Journal of Child Language* 2: 65–77.
MacWhinney, Brian
1976 Hungarian research on the acquisition of morphology and syntax. *Journal of Child Language* 3: 397–410.
MacWhinney, Brian
1978 The acquisition of morphophonology. *Monographs of the Society for Research in Child Development* 43 (whole no. 1): 1–123.
MacWhinney, Brian
1982 Basic syntactic processes. In *Language Acquisition: Vol. 1. Syntax and Semantics*, Stan Kuczaj (ed.), 73–136. Hillsdale, NJ: Lawrence Erlbaum.
MacWhinney, Brian
1987 The Competition Model. In *Mechanisms of Language Acquisition*, Brian MacWhinney (ed.), 249–308. Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, Brian
2000 *The CHILDES Project: Tools for Analyzing Talk*. 3d ed. Mahwah, NJ: Lawrence Erlbaum Associates.
MacWhinney, Brian
2004 A multiple process solution to the logical problem of language acquisition. *Journal of Child Language* 31: 883–914.
MacWhinney, Brian
2008a Cognitive precursors to language. In *The Evolution of Communicative Flexibility*, Kimbrough Oller and Ulrike Griebel (eds.), 193–214. Cambridge, MA: MIT Press.
MacWhinney, Brian
2008b Enriching CHILDES for morphosyntactic analysis. In *Trends in Corpus Research: Finding Structure in Data*, Heike Behrens (ed.), 165–198. Amsterdam: John Benjamins.
MacWhinney, Brian
2008c How mental models encode embodied linguistic perspectives. In *Embodiment, Ego-Space, and Action*, Roberta Klatzky, Brian MacWhinney and Marlene Behrmann (eds.), 369–410. Mahwah: Lawrence Erlbaum.
MacWhinney, Brian
2009 The emergence of linguistic complexity. In *Linguistic Complexity*, Tom Givon and Masayoshi Shibatani (eds.), 405–432. New York: Benjamins.
MacWhinney, Brian
2012 The logic of the Unified Model. In *The Routledge Handbook of Second Language Acquisition*, Susan Gass and Alison Mackey (eds.), 211–227. New York: Routledge.
McDonald, Janet L.
1986 The development of sentence comprehension strategies in English and Dutch. *Journal of Experimental Child Psychology* 41: 317–335.
McNeill, David
1979 *The Conceptual Basis of Language*. Hillsdale, NJ: Lawrence Erlbaum.
Mel’cuk, Igor, and Alain Polguère
1987 A formal lexicon in the Meaning-Text Theory: Or how to do lexica with words. *Computational Linguistics* 13: 1–88.
Menyuk, Paula
1969 *Sentences Children Use*. Cambridge, Mass.: MIT Press.
Monaghan, Padraic, and Morten Christiansen
2010 Words in puddles of sound: Modelling psycholinguistic effects in speech segmentation. *Journal of Child Language* 37: 545–564.

Mooney, Raymond
2008 Learning to connect language and perception. Paper presented at the 23rd AAAI Conference on Artificial Intelligence, Chicago, IL.

Nelson, Katherine
1973 Structure and strategy in learning how to talk. *Monographs of the Society for Research in Child Development* 38, whole Nos. 1–2.

O'Grady, William
2005 *Syntactic Carpentry*. Mahwah, NJ: Lawrence Erlbaum Associates.

Perfors, Amy, Joshua B. Tenenbaum, and Elizabeth Wonnacott
2010 Variability, negative evidence, and the acquisition of verb argument constructions. *Journal of Child Language* 37: 607–642.

Pienemann, Manfred, Bruno Di Biase, Satomi Kawaguchi, and Gisela Håkansson
2005 Processing constraints on L1 transfer. In *Handbook of Bilingualism: Psycholinguistic Approaches*, Judith F. Kroll and Annette M. B. DeGroot (eds.), 128–153. New York: Oxford University Press.

Pine, Julian M., and Elena V. M. Lieven
1997 Slot and frame patterns in the development of the determiner category. *Applied Psycholinguistics* 18: 123–138.

Pinker, Steven
1982 A theory of the acquisition of lexical-interpretive grammars. In *The Mental Representation of Grammatical Relations*, Joan Bresnan (ed.), 655–726. Cambridge, MA: MIT Press.

Pinker, Steven
1995 Why the child holded the baby rabbits: A case study in language acquisition. In *An Invitation to Cognitive Science*, vol. 1: *Language*, 2d ed., Lila Gleitman and Mark Liberman (eds.), 107–137. Cambridge, MA: MIT Press.

Poeppel, David, and Kenneth Wexler
1993 The full competence hypothesis of clause structure in early German. *Language* 69: 1–33.

Pollard, Carl, and Ivan Sag
1994 *Head-Driven Phrase Structure Grammar*. Chicago: Chicago University Press.

Pulvermüller, Friedemann
2003 *The Neuroscience of Language*. Cambridge: Cambridge University Press.

Rowland, Caroline F., and Julian M. Pine
2000 Subject-auxiliary inversion errors and wh-question acquisition: "What children do know?" *Journal of Child Language* 27: 157–181.

Roy, Deb
2002 Learning visually grounded words and syntax for a scene description task. *Computer Speech and Language* 16: 353–385.

Rytting, C. Anton, Chris Brew, and Eric Fosler-Lussier
2010 Segmenting words from natural speech: Subsegmental variation in segmental cues. *Journal of Child Language* 37: 513–543.
Sagae, Kenji, Eric Davis, Alon Lavie, Brian MacWhinney, and Shuly Wintner
2007 High-accuracy annotation and parsing of CHILDES transcripts. *Proceedings of the 45th Meeting of the Association for Computational Linguistics*. 1044–1050. Prague: ACL.
Sagae, Kenji, Eric Davis, Alon Lavie, Brian MacWhinney, and Shuly Wintner
2010 Morphosyntactic annotation of CHILDES transcripts. *Journal of Child Language* 37: 705–729.
Sagae, Kenji, Alon Lavie, and Brian MacWhinney
2005 Automatic measurement of syntactic development in child language. *Proceedings of the 43rd Meeting of the Association for Computational Linguistics*: 197–204. Ann Arbor: ACL.
Sagae, Kenji, Brian MacWhinney, and Alon Lavie
2004a Adding syntactic annotations to transcripts of parent-child dialogs. *Proceedings of the Fourth International Conference on Language Resources and Evaluation LREC* 2004: 1815–1818. Lisbon: LREC.
Sagae, Kenji, Brian MacWhinney, and Alon Lavie
2004b Automatic parsing of parent-child interactions. *Behavior Research Methods, Instruments, and Computers* 36: 113–126.
Schlesinger, Izchak M.
1974 Relational concepts underlying language. In *Language Perspectives: Acquisition, Retardation, and Intervention*, Richard L. Schiefelbusch, and Lyle L. Lloyd (eds.), 129–151. Baltimore: University Park Press.
Simon, Herbert
1969 *Sciences of the Artificial*. Cambridge, MA: MIT Press.
Siskind, Jeffrey
2000 Learning word-to-meaning mappings. In *Models of Language Acquisition*, Peter Broeder, and Jaap Murre (eds.), 121–153. Oxford: Oxford University Press.
Sokolov, Jeffrey. L.
1993 A local contingency analysis of the fine-tuning hypothesis. *Developmental Psychology* 29: 1008–1023.
Steedman, Mark
2000 *The Syntactic Process*. Cambridge, MA: MIT Press.
Stemberger, Joseph Paul
1985 Bound morpheme loss errors in normal and agrammatic speech: One mechanism or two? *Brain and Language* 25: 246–256.

Tenenbaum, Joshua B., Thomas L. Griffiths, and Charles Kemp
2006 Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences* 10: 309–318.
Terrace, Herbert S., Laura-Ann Petitto, Richard J. Sanders, and Thomas Bever
1980 On the grammatical capacity of apes. In *Children's Language*, vol. 2, K. Nelson (ed.), 371–496. New York: Gardner Press.
Tesniére, Lucien
1959 *Elements de syntaxe structurale*. Paris: Klincksieck.
Thiessen, Erik D., and Jenny R. Saffran
2007 Learning to learn: Acquisition of stress-based strategies for word segmentation. *Language Learning and Development* 3: 75–102.
Tomasello, Michael
1992 *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.
Tomasello, Michael
2000 The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences* 4: 156–163.
Tomasello, Michael, and Daniel Stahl
2004 Sampling children's spontaneous speech: How much is enough? *Journal of Child Language* 31: 101–121.
Vosse, Theo, and Gerard Kempen
2000 Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition* 75: 105–143.
Waterfall, Heidi, Ben Sandbank, Luca Onnis, and Shimon Edelman
2010 An empirical generative framework for computational modeling of language acquisition. *Journal of Child Language* 37: 671–703.
Wexler, Kenneth, and Peter Culicover
1980 *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.
Worden, Robert.
2002 Linguistic structure and the evolution of words. In *Linguistic Evolution through Language Acquisition*, Ted Briscoe (ed.), 75–110. Cambridge: Cambridge University Press.
Zelle, John, and Raymond Mooney
1993 Learning semantic grammars with constructive inductive logic programming. Paper presented at the AAAI – 93.