Handbook of Categorization in Cognitive Science, ed. by Henri Cohen & Claire Lefebvre, 2005, Elsevier. 433-457

The acquisition of grammatical categories : a state of the art

Marie Labelle, Université du Québec à Montréal

Abstract

This contribution reviews recent research on the acquisition of grammatical categories, focusing on three aspects of the problem: parts of speech, inflection, and subcategories of words. It appears that children are sensitive to the distribution of the various classes of words surprisingly early, and that inflectional categories are mastered by age three despite the complexity of their acquisition. Subcategories of words take longer to acquire. To account for the facts, it is generally assumed that children are equipped with a distributional learning mechanism, but given the complexity of the factors that must be taken into account, some constraints must be placed on this mechanism. Various types of possible constraints are discussed, from innate knowledge to processing constraints.

1. Grammatical categories

The term “grammatical category” covers a variety of types of categories. We focus here on parts of speech, word classes, and inflectional categories. *Part of speech* refers to the familiar categories of words like noun, verb, adjective. In many cases, the words of these general categories split into subcategories having partially distinct grammatical properties. We will call these subcategories *word classes.* In addition, languages vary as to the grammatical or *inflectional* markers that appear on words. (See Table 1 for examples).

(insert Table 1 about here)

Why is it important to know the grammatical category of a word? What does it mean to say that a child "knows" that some element is a 'Noun' ? The answer to the first question is that knowing a word's category is a precondition for knowing how to use the word in the language. The grammatical category of a word determines (1) the position it is allowed to occupy in the clause (German verbs appear in final position in subordinate clauses and in second position in main clauses); (2) the range of syntactic functions it can occupy (a noun may be the subject of a clause, but a preposition or an adjective cannot); (3) the type of words with which it co-occurs (determiners co-occur with nouns, but not with verbs); (4) the types of morphemes it requires or accepts (verbs inflect for tense; nouns inflect for number...) How do we know that the missing word in sentence (1) is a noun? It is not because we can look it up in the dictionary, but because it occupies the typical position of a noun, is modified by an adjective, is preceded by a determiner and so on. Similarly, it is on the basis of grammatical properties that we know that the word *butter* in (2) is a verb and not a noun.

(1) The blue ___ is on the table.

(2) The children *butter* their toasts.

Thus, to say that a child knows that some word is a 'Noun' is to say that he or she uses this word in the typical positions occupied by nouns, with the proper inflection, and so on. To produce grammatical sentences in the language they are acquiring, children must use the words according to the properties determined by their category. The specific label attached to a part of speech is not what interests us here. What matters is the fact that the different parts of speech have different grammatical properties. This is

clearly expressed by Pinker (1984 :43) : “The only significance of the name of a symbol (like Noun) in a cognitive account is that the process that manipulates such symbols treats all the symbols with a given name alike but differently from symbols with different names.”

The child learning a language must categorize every new word according to its part of speech, and in some cases according to some particular word class within the part of speech. The child must also identify the inflection occurring on words and categorize it as denoting gender, number, grammatical role or any of the possible inflectional categories. Moreover he or she must assign each word to the proper inflectional class, as it is frequently the case that the phonological form of inflectional markers is not uniform across the language. For example, the ending for Nominative plural in German is not the same for all masculine nouns, as illustrated in Table 2.

(insert Table 2 about here)

We will start with parts of speech, then move on to inflectional categories and finally to word classes.

2. Two-word utterances and their analysis

Let us consider some typical two-word combinations produced by English-speaking children around their second birthday, and some characterizations of them, displayed in Table 3. The grammatical description in the second column assumes that young children categorize words into the part of speech categories of the adult language. Under the semantic approach typical of the seventies (third column), children’s first word

combinations are the result of rules mappings over semantic categories (Bowerman 1973, Schlesinger 1971). In Braine's Pivot grammar (Braine 1963), fourth column, the grammar of young children contains only two categories of words : Pivot words and Open words. The latter two approaches must explain how and when children abandon the initial grammar and move on to a grammar based on part of speech categories, recategorizing in the process every word in terms of the adult grammar. As we will see, recent evidence shows that, from the beginning, children attend to the formal properties of words.

(insert Table 3 about here)

Table 3 also illustrates the fact that the utterances produced by English-speaking children around 24 months are devoid of functional elements like determiners, auxiliaries, and morphological endings. Some authors have suggested that children might not attend these elements, presumably because they are unstressed (Echols & Newport 1992). Their internalized grammar would not yet contain functional categories (Radford 1995, 1996). Here again, we will see that children take functional elements into account long before they display knowledge of these categories in their production.

3. A semantic approach to grammatical categorization: Semantic bootstrapping

How does a child learn to categorize words as being verbs, nouns, adjectives, or adverbs? A view that came to be known under the term **semantic bootstrapping** is based on the idea that semantics guides the child into categorizing the words of his language: a word denoting an object or a person is a noun, a word denoting an event is

a verb, and so on. Children would use semantics to "boot" their way into syntax. There are two main approaches to semantic bootstrapping.

A **cognitive approach** assumes that infants organize the world into cognitive categories like objects, events, and so on. When they start learning language, they initially classify the words they learn in terms of these cognitive categories (Brown 1973, Macnamara 1982). Grammatical categories like 'noun' or 'verb' emerge when children start learning the properties of words.

The **innatist hypothesis** (Grimshaw 1981, Pinker 1984) assumes that the child is innately equipped (1) with two types of universal knowledge, a set of part-of-speech categories and a set of semantic categories, and (2) with a universal default mapping between the two, as shown in Table 4. The child infers the categorization of words on the basis of their semantic properties.

(insert Table 4 about here)

Semantic criteria do not provide an adequate basis for determining the part of speech of a word. Many nouns do not refer to physical objects, some denote actions (*a hug*), events (*an explosion*), or states (*a depression*); many verbs do not refer to actions (*want*), some denote a property (*resemble*); some adjectives are actional (*quick, fast*) or controllable (*be polite !*) or denote concepts expressed in other languages with verbs (*hungry, wet, sleepy*), and so on (Maratsos and Chalkley 1980, Maratsos 1999 : 206). If children start out with semantic categorization criteria, at some point they must abandon them, otherwise they will not converge on the proper grammar. In both approaches to semantic bootstrapping, only the first set of words in each category is semantically

classified. Once some elements have been classified, the child starts analyzing the grammatical properties of these words, and uses the distribution of the items which do not conform to the basic semantic type to posit their category on the basis of known structure. Gradually, grammatical properties come to predominate as classifying devices (Macnamara 1982).

Semantic bootstrapping came under criticism both on theoretical and on empirical grounds. First, Benedict (1979) classified the first 50 words comprehended and produced by eight children according to semantic criteria based on the child's use of these words. Her class of 'action words' contained interjections (*peekaboo*), verbs (*eat, give*), adverbs (*no*), particles (*out, up, down*), and her class of 'modifiers' contained adjectives (*big, hot*), adverbs (*there*), pronouns (*mine*), and unanalyzed phrases (*allgone*). If children classified all these words as verbs in the first case, and as adjectives in the second, they would be in deep trouble. Yet, there is no evidence that children miscategorize words to a large extent (Maratsos and Chalkley 1980, Braine 1987). This suggests that children attend from the beginning to distributional information.

Second, semantic bootstrapping would be supported if it could be shown that the first nouns that children learn designate objects or people and their first verbs, actions. But Bassano (2000), in a study of the first words produced by a French-speaking child, observes that concrete action verbs are not the earliest verbs to be produced with any frequency; situational, attention-getting verbs, modals and *être* 'be' are the first to occur. She also reports that neither nonconcrete object names nor nonconcrete action verbs are avoided in the child's early speech.

Third the innatist approach to semantic bootstrapping crucially assumes that there is a universal set of grammatical categories. Many typologists would reject this

assumption (Culicover 1999 :39). For example, in Japanese, there are two categories of adjectives (two grammatically distinct sets of words denoting properties), while in other languages there are no adjectives, and adjectival meanings are expressed by nouns (words denoting properties are non-distinct grammatically from words denoting entities) (3) or by verbs (4) (Schachter 1985 :32-33).

(3) a. Rikashka : hatun-kuna-ta (Quechua)
I-saw big-pl-acc
'I saw the big ones.'

b. Rikashka : alkalde-kuna-ta (Quechua)
I-saw mayor-pl-acc
'I saw the mayors.'

(4) a. Piaoliang de nühaizi (Mandarin Chinese)
beautiful rel girl
'a girl who is beautiful, a beautiful girl'

b. Liaojiede nühaizi (Mandarin Chinese)
understand rel girl
'a girl who understands, an understanding girl'

Even such a basic distinction as that between Noun and Verb does not appear to be universal. In **multicategorial** languages like Tagalog, entity-denoting roots and event-denoting roots seem to have identical grammatical properties; both may function as predicates or as arguments (Schachter 1985 :11).

(5) a. Nagtatrabaho ang lalaki (Tagalog)
is-working TOP man
'The man is working.'

b. Lalaki ang nagtatrabaho
man TOP is-working
'The one who is working is a man.'

Multicategoriality poses a major problem to a semantic bootstrapping approach. This is seen more clearly if we consider the multicategoriality of some frequent English words: *kiss* is a noun in *Give me a kiss* but a verb in *Kiss me* (also *hug, drink, bite, call*, etc.) (Nelson 1995). If children interpret *kiss* as an action word, and assume that it is a verb, and if they hear *kiss* sometimes in a verb frame and sometimes in a noun frame, how are they to infer the grammatical properties of verbs and nouns? In languages like Tagalog, where multicategoriality is generalized to the whole lexicon, the syntax of the language cannot be learned via semantic bootstrapping.

We conclude that, while semantic categories certainly participate in word classification in that if a child notices that a word denoting an action has property x, he or she might expect other words denoting an action to have the same property, these categories do not form the initial basis for grammatical categorization. To converge on the adult grammar, children must register the distributional properties of lexical items.

4. Distributional learning

A number of computer simulations have shown that distributional learning can go a long way into categorizing the words of the language on the basis of positional and co-

occurrence information (e.g. Cartwright and Brent 1997 ; Brent 1994, 1996). Also, studies of caretakers' language show that distributional information is a reliable cue to grammatical categorizes in the language children hear (e.g. Mintz et al. 2002). Moreover, psychological experiments with artificial languages demonstrate that not only adult but also infant learners are able to perform distributional analyses (e.g. Braine et al. 1990, Brooks et al. 1993, Mintz 2002, Saffran, Aslin and Newport 1996, Valian and Coulson 1988). Infants have thus the **capacities** and the **proper input** to learn grammatical categories on the basis of distributional information. In this section, we discuss evidence showing that children register formal cues in the language and use them to categorize words. We distinguish three types of formal cues: word order, morphological cues, and co-occurrence restrictions.

4.1. Word order

Studies of **comprehension** indicate an early sensitivity to word order. Using a preferential looking paradigm, Hirsch-Pasek and Golinkoff (1996) show that infants as young as 16 months distinguish "Big Bird is tickling Cookie Monster" from "Cookie Monster is tickling Big Bird": upon hearing the first sentence, they tend to look longer at a video showing Big Bird doing the action (over a video where Cooking Monster is doing the action).

In language **production,** the earliest moment when we can start observing effects of word order is when children start putting two words together. As long as children produce one word at a time, we can't talk of word order. Do the first two-word utterances produced by young children already show an effect of word order? Yes. Children produce utterances like: *more juice, more car, more read, more hot*, with *more* in first

position, and *boot off, hat off, shoe off*, with *off* in second position. These two-word utterances seem to be the product of highly restricted **limited-scope formulae** containing a relational lexical item (*more, off*) and a positionally specified empty slot for its argument: more+X, X+off (Braine 1976). The word order consistency shows that children register recurring combinations in the input, and are attentive to positional cues.

Children seem to **treat nouns and verbs differently** almost from the beginning. In English, many verbs alternate between a transitive use and an intransitive use where the subject of the intransitive corresponds to the object of the transitive :

(6) a. John broke the vase.

b. The vase broke.

The alternation can be described as a rule allowing a verb to switch between two argument frames. Tomasello et al. (1997) taught novel verbs and novel nouns (e.g. *wug*) to children aged between 1;6 and 1;11. The words were taught in one frame, and the authors looked at whether the children could generalize to other frames. They found that children were conservative with novel verbs, which they tended to use only in the type of frame in which they were taught, reproducing the modeled word order. But when taught a novel noun, they could instantly use it in the two-word combinations that they mastered, either in first or in second position (*wug gone, more wug*). The authors suggest that children of that age possess some kind of paradigmatic category “Noun” corresponding to the type of item that can fill the variable slot in the limited-scope formulae that they know, but not of the category “Verb” (Tomasello & Brooks 1999 : 168). Fisher (2002) argues that the difference in treatment of nouns and verbs is due to

the different syntactic role that verbs play in clauses: verbs are **relational** elements. Naigles (2002, 2003) also points out that children may be reluctant to generalize verbs across frames because a change of frame implies a change of perspective on the event, whereas this is not the case with nouns. In comprehension, children seem to be better able to switch from one frame to another (Naigles and Hoff-Ginsberg 1995).

This research shows that children distinguish relational terms from non-relational terms. When learning names for entities, they encode only the type of referent, but when learning relational terms like verbs or words like *more, off*, they encode both the approximate meaning of the term and its argument frame, more or less as in (7).

(7) *more* : meaning : recurrence of X
argument frame : *more*+X

In (7), the position of the argument is specified, but not its grammatical category ("Noun"). Children produce not only utterances like *more cookie*, with a Noun as complement of *more,* but also *more read* 'read some more' or *more hot* 'another hot thing' (Braine 1976). This shows that the fully specified grammar of column 2 in Table 3 attributes too much knowledge to the child.[1] In the early grammar, the argument frame is probably underspecified as to the grammatical category of the complement.

The absence of generalization with verbs ties in with other research showing that children learn constructions lexical item by lexical item and do not generalize from one

[1] For the same reason, the open slot is not specified in semantic terms, like *more+substance*, contrary to the semantic approach illustrated in Table 3.

lexical item to another (Braine 1976, Peters 1983, Tomasello 1992). Indeed Akhtar and Tomasello (1997) have shown that when children younger than 3;0 hear a novel transitive verb for which the argument structure has not been modeled ("This is called *dacking*"), they are unable to use nor comprehend word order to mark agents and patients. One is reminded of Karmiloff-Smith's model (1986) according to which the initial phase of learning is a **list of independent procedures**; these procedures are integrated into more general ones at a later stage involving representational redescription.

The conceptual simplicity of nouns as labels for objects compared to the more complex role of verbs as relational elements has been proposed (Gentner 1982) to account for the predominance and earlier acquisition of nouns over verbs in English, Italian, Spanish, and French (Bates et al. 1994, Caselli & al. 1999, Jackson-Maldonado et al. 1993, Bassano 2000). This is the so-called "noun bias" in acquisition. As the noun bias is not observed in the acquisition of Asian languages (Choi 1997, Choi and Gopnik 1995, Gopnik & al. 1996), it seems that language-specific factors are at play; for example, Asian languages allow massive argument deletion, resulting in many utterances containing only a verb (de Boysson-Bardies 1996).

4.2. Inflection and inflectional class

Much of the information necessary to determine the grammatical category of a word is coded by **function words** like determiners or auxiliaries, and by **inflectional morphemes**. Thus, upon hearing the nonsense clause *The wug zaks*, we know that *wug* is a noun (because it is preceded by *the*) and *zaks* a verb (because it agrees with *wug*). Valian and Coulson (1988) have shown that frequent markers are anchor points

for distributional analysis: adults have more success in learning an artificial language when the language contains frequent markers. When do children start registering function words and inflection? Do they use this information to categorize words? We might hypothesize for example that a child learning English would observe *-ed* on certain words and use this cue to determine membership in the category Verb. While this appears trivial, it is not. Affixes (prefixes or suffixes) can be inflectional, but they can also be category-changing: adding *–ly* to adjectives turns them into adverbs. As stressed by Culicover (1999 :38): "The formal type of knowledge can only be brought into play after the basic categories are established, at least in some preliminary way. Once a learner knows that a particular formal inflection is a property of a certain category, this information can be used to categorize new elements." In order to use some affix as a cue to the category of a word, the child must first **identify** the affix and second determine on which **type of word** it appears. But an inflectional morpheme is also a grammatical element that must be **categorized** as **marking some grammatical function**. In the present section, we focus on inflectional morphology as a categorization problem; in section 4.3 we turn to the question of whether children use function words and inflectional morphology to determine the category of words.

Categorizing an inflectional morpheme as marking some grammatical function is not a simple matter, as the marker-to-function mapping is often many-to-one or one-to-many. A case of **one function/many forms** is observed in declension classes. In inflected languages, it is frequently the case that a part-of-speech category is subdivided into subclasses, each one accepting a specific set of forms (see Table 2 for German masculine nouns in the nominative plural). In Latin there were five declension classes. In Bantu languages, there may be up to 20 different word classes indicated by obligatory

prefixes (Suzman 1996). Such classes are in general arbitrary without any semantic core to them (Braine 1987). In addition, in many cases, the pronunciation of an affix varies according to its phonological environment, and the child must determine that the various forms he hears are grammatical variants of the same underlying morpheme (Peters 1997). In some cases, the phonological changes brought about by a morpheme are so complex that it is difficult to identify the derived word as being fundamentally "the same word" as the initial one. (In Chickasaw, *hilhali* 'I'm dancing' becomes in the negative *akhi'lho* 'I'm not dancing'—Anderson 1985 : 165)

Homonymy leads to a situation of **one marker/many functions**. To give some examples, in Polish nouns, the suffix *-a*, which marks nominative case in class 1 nouns, marks genitive case in class 2 nouns (Maratsos 1999). The clue provided by the affix can be relied on only after the learner has some knowledge of the declension classes. In English, *-s* is not a reliable cue to the category Verb, as it marks verbs for third person singular, nouns for plural, and it is also a possessive marker.

An example provided by Maratsos (1998, 1999) is worth citing in full. In Turkish, functions are marked by inflectional case markers. The suffix *-u* on a noun means that it is the patient of the action (the direct object). How can the child 'notice' that -u marks patienthood? As word order is quite free, the child cannot rely on it. Suppose the child hears the Turkish equivalent of "Sam-u scratched Ann" and can tell from the context that Sam is the patient. Is it sufficient to learn that -u marks patienthood? No, says Maratsos, because -u could mark many other things: gender (masculine), the word-class of Sam (long, thin object), humanness, animacy, third person, social status, affectionate regard on the part of the speaker, or any other category found on nouns in languages. To determine the value of the suffix, the child must be able to register the properties

actually encoded by the language. If we add to that the fact that, in Turkish, not all patients are marked with -u, only definite patients, and that there are actually four variants of the -u suffix (/u/, /ü/, /i/ and /ï/) chosen according to the phonological characteristics of the stem, so that no single form consistently marks patients, we can see that learning the value of the sound /u/ is not an easy matter, because this form appears on a small subset of patients, perhaps 10% to 15%. A serial induction process in which the child makes a guess at the property encoded by the form and tests the validity of this guess against the input, one at a time, until one is found that works (Pinker 1984) would have to have a very low threshold of success to be applicable. On the other hand, if the child registers all possibilities at once and then checks off the ones that don't match upon subsequent exposures to the suffix, he or she would have to consider not only the possible values of the form, but also the phonological characteristics of the stem and the definiteness of the patient. Maratsos concludes that the child can't just 'notice' the value of inflectional morphology upon a single exposure; he or she must 'grind through' a large number of possibilities over a wide variety of utterances.

If we add to that the fact that a single affix may convey more than one piece of information (the ending *–ai* on the French verb *marchai* 'walked' marks both tense—Passé Simple—and person-number of the subject —first person singular); and that in highly inflected languages like Greenlandic Eskimo (8), roots are constructed with more than one affix, each one requiring identification and categorization, we have to conclude that learning inflection is not an easy matter.

(8) uppi-ti- le- qa- akkit

fall- cause-begin-intensifier-1st/2nd singular indicative

"I'm going to make you fall!"

How do children cope with such difficulties? Surprisingly perhaps, cross-linguistic studies show that children make very few errors. When they use inflectional morphology, they use it right; errors are more often of omission than of commission (Maratsos 1998; Phillips 1995). Moreover, children master the essentials of the inflectional system of their language by their third birthday provided the system is regular and phonologically transparent. The case system of Turkish and Polish is mastered with no observed error by age two (Aksu-Koc and Slobin 1985, Smoczynska 1985). Children learning highly inflected languages produce inflectional markers productively much earlier than English-speaking children; example (8) was produced by a two year old (Fortescue & Lennert Olsen 1992) (see also Choi 1997, Clancy 1985). The complex word-class system of Bantu nouns seems to be mastered with almost 100% accuracy by 2;6 (Suzman 1996). When inflectional paradigms are phonologically complex, riddled with homophony, semantically opaque, or inconsistent, children master the essential of the system by their third birthday, but make errors with exceptional forms, which they tend to regularize.

The robustness and efficiency of the acquisition of inflectional morphology suggest that some innate factors are at work, faculty-specific, species-specific, or both (Maratsos 1998). As the conceptual domains which are grammaticalized in languages seem to form a closed list, it is tempting to think that learning would be helped if children limited their search to that list. The child might also be helped by an innate knowledge of the universal correspondence between some grammatical categories and some types of

markers (see Table 5). Once the child has identified a suffix indicating tense, he would automatically know that the word it is suffixed to is a verb.

(insert Table 5 about here)

But this strategy would not work in every case. In Nootka, the inflectional markers for tense and mood attach to the right of the first word in the sentence; in (9) this word happens to be the accusative case marker of the noun phrase *the deer* 'ʔooqw bowatc ʔaq' (Anderson 1985 : 156, ex. 1a-b). In Kwakw'ala (10), grammatical particles attach phonologically to the **preceding** word rather than to the word they determine (Anderson 1985 :166, ex. (2)).

(9)	ʔooqw-**obt-ʔa**	bowatc	ʔaq	tl'itcitl-	John	
	acc **-past-declar**	deer	det	shoot	John	
	'The deer, John shot (it).'					
(10)	Nep'id-*i-da*	gʔnanʔm-*xa*	guk^{w}sa	t'isʔm		
	throw-SUBJ-DET	child-OBJ	house-INST	rock		
	'The child threw a rock at the house.'					

While innate knowledge constraining the search space would help, we have to concur with Maratsos that children must grind through the data to discover the correct generalizations.

As to the idea dear to cognitivists that the acquisition of grammatical markers is driven by the need on the part of the child to express some cognitive function, Bowerman (1985) argues that this can't be the case. The acquisition of determiners in French or English is not driven by the child's desire to communicate whether a noun is already known to the speaker—definite (*the dog*)—or not—indefinite (*a dog*), but by the necessity, grammaticized in his language, of encoding it. We are reluctant to recognize this because familiarity makes us take for granted the distinctions encoded in our own language, but, says Bowerman, "consider the obligatory four-way classification of nouns in sentences of Toba, a language of Argentina, according to whether the objects to which they refer are in view, out of view, coming into view, or going out of view, and furthermore, if they are in view, according to whether they are spatially nonextended (e.g. a fruit), extended vertically (e.g., a fruit still hanging, or a tree), or extended horizontally (e.g., a table)(Klein, 1979). Can such meanings really struggle for expression in the developing minds of all children, including those in our own living rooms ?" (Bowerman 1985 :377).

In fact, one recurrent observation in the language acquisition literature is that formal distinctions orient the child towards discovering the semantic relations encoded by them (rather than cognitive distinctions orienting the child towards finding the formal way to express them) (Bowerman and Choi 2001, Slobin 2001). Phonology, in particular, plays a crucial role in the acquisition of morphology. Children learning Bantu languages rely, not on semantics, but on phonological information to learn the complex noun class system (Demuth 1992; Suzman 1996). Demuth suggests that this holds crosslinguistically : "Access to the semantics of the system becomes available only at later stages of development, whereas early overgeneralizations are normally of a

phonological nature" (Demuth 1992 :630) . The same point is made by Naigles (2002), who argues that **children find form easy, but pairing of form and meaning hard**, in essence because it requires more computational resources (see also Peters 1997).

If children must discover the function of morphemes by observing the way they are used in the input, we might expect them to entertain erroneous hypotheses for a while. This is discussed by Eve Clark (2001) under the term 'emergent categories'. Clark suggests that children may for a brief while use some morpheme of their language to express distinctions made in other languages, but not in the language learned. She gives the example of a child who encoded the difference between inherent properties and temporary properties by using respectively adjectives in -y (*It's crumby* = full of crumbs, said of a biscuit ; D 2 ;6,9) and adjectives in -ed (*My foot is all crumbed* = covered with crumbs ; D 2 ;6 ;30.). Again, this raises the question : are there built-in constraints as to the kinds of hypotheses children might entertain ?

4.3. Function words

We have seen that children attend to inflectional morphology from the beginning. Another cue to the category of a word is provided by the types of function words with which it co-occurs. Do children register the function words in sentences, and do they use them to determine the grammatical category of content words? Children learning English might for example notice that words denoting concrete objects are preceded by a determiner like *a* or *the*. Even though not every concrete object is immediately preceded by a determiner (*a big truck*) and not every word preceded by a determiner denotes a concrete object (*a hug*), they could use the statistical regularity to infer that a novel word preceded by a determiner (*a wug*) might denote a concrete object.

In a seminal study, Brown (1957, replicated in Dockrell & McShane 1990) taught a novel word to sixteen children of three to five years of age, while presenting them with pictures of some novel activity performed on some confetti-like material. Three different formulations were used. In the verb frame, the question was : *"Do you know what it means to sib ? In this picture, you can see sibbing. Now show me another picture of sibbing."* In the count noun frame, the question was : *"Do you know that a sib is ?"*, and in the mass noun frame, it was : *"Have you ever seen any sib ?"* When the new word was introduced as a verb, the children tended to choose the picture displaying the same action (10/16); with the count noun frame, they chose the picture displaying the same object (11/16); and with the mass noun frame, they chose the similar substance (12/16). Thus, three to five year olds infer the part of speech of a word on the basis of its structural environment and use this information to determine the meaning of words.

Function words—*functors*—like determiners and auxiliaries are not only cues to the grammatical category of other words; they are themselves grammatical elements which must be categorized. They are characterized by a cluster of phonological properties like lack of stress, monosyllabicity, short syllable duration, null coda and vowel harmony, which distinguish them from full lexical items (Shi, Morgan & Allopenna 1998), but which also make them less salient. When do children start taking these elements into account? Within the generative framework, functional elements head syntactic phrases, building a functional structure above lexical items. A central theoretical problem is exactly how much functional structure children master. The hypotheses range from a **full competence** approach, all functional categories assumed to be present from the start (Borer and Rohrbacher 1997, Poeppel & Wexler 1993), to **structure building** whereby the functional structure is built up as children acquire the

functional elements (Radford 1995,1996) (For discussion see Guasti 2002.) In the latter case, the appearance of filler phonemes (the proto-determiners /i/, /a/, /n/ in (11)) where a functional morpheme is expected would denote the beginning of an awareness of the presence of the corresponding functional category (Peters 1995:472):

(11)	/i ká:/	'F car'	(F = filler)
	/a gÚdki/	'F cookie'	
	/n báp/	'F bump'	

The structure building hypothesis seems to account for the gradual accretion of functional markers in language production, but it is confronted with the problem of explaining the difference between comprehension and production. The child who does not produce some functional morpheme should not be able to use it in comprehension as a cue to the grammatical category of words. But a growing number of studies shows that functional elements are taken into account long before children start producing them. For example, 24-month old single word talkers react negatively to sentences containing incorrectly used grammatical morphemes or where the grammatical morpheme is omitted (Golinkoff, Hirsch-Pasek & Schweisguth 2001). At a stage when they do not produce grammatical morphemes, children carry out fewer commands when the grammatical morphemes are omitted (*Throw ball* vs *Throw the ball*) (Shipley, Smith, and Gleitman 1969). When asked to imitate sentences containing real or pseudo-functors, children with a mean length of utterance (MLU) lower than 2 words leave out the real morphemes more often than the pseudo morphemes (Gerken, Landau & Remez 1990). Children confronted to grammatical stimuli of the type *Find <u>the</u> bird for me* and

ungrammatical stimuli with a pseudo-functor *Find gub bird for me* or an erroneous functor *Find was bird for me* respond correctly more often to the grammatical stimuli, showing that they are aware of where the real functors occur and what they sound like (Gerken & McIntosh 1993). Even children with MLU less than 1.5 who spontaneously produced no determiner at all responded significantly better to commands with grammatical morphemes. Already by 16 months, children have an idea of where function words are likely to occur, as they show listening preference for natural passages, over passages where function words have been misplaced (Jusczyk 2001). Finally, infants as young as 10.5 months distinguish normal English passages from passages where nonsense syllables replace function words (Shady 1996).

Such studies show that very young children notice grammatical elements and are sensitive to their distribution. But are they able to use morphology and function words to determine a word's part of speech? Waxman showed that when presented with a novel word in a noun frame (*These are blickets ; this one is a blicket*) or in an adjective frame (*These are blickish ; this one is blickish*), children of 14 months with a mean production vocabulary of 15 words map nouns to categories of objects (Waxman and Booth 2001). They attend to the cue provided by the determiner or the plural morpheme up to a year before displaying these elements in their production. The adjective frame did not consistently elicit a property interpretation at 14 months, but preliminary evidence suggests that it did by 24 months.[2]

---

[2] The main cue for the property interpretation in the adjective frame used is the infrequent ending *-ish* on the word, and the syntactic position of *blickish* after the copula

To sum up, very young children are aware of the morphology appearing on lexical items and of the position and form of function words, and they are able to use this knowledge to determine the grammatical category of words and their interpretation. Just how much knowledge of function words children have at this early stage of language learning is a question that we are far from being able to answer at this point.

Why is it that children seem to understand so much, yet produce so little? It could be that children don't know enough about the function or meaning of functional morphemes to be able to integrate them in their productive grammar. For example children might have categorized *the, a, an* as belonging to the category D of determiners, and might use this knowledge to determine that the following word is a noun, while not having attached features like [+definite], [+singular] or [before Vowel] to them. Since such features determine which determiner to chose in a given context, children don't have enough knowledge of these words to use them. Observe that the proto-determiner /n/ in (11), presumably linked to *an*, is used in front of a word beginning with a consonant. The child has an idea of what determiners sound like but he doesn't quite know which form to chose in a given context.

Gerken (2001 and references therein) argues that the omission of grammatical morphemes in the language of young children is to be attributed to a metrical constraint on production. Children's utterances would tend to be limited to sequences of weak syllables alternating with strong syllables, and grammatical morphemes would be omitted when they don't fit into this pattern. Children are likely to say *He hugs the dog*

---

in *This one is blickish* is also that of proper names (*This one is Fred*), and of mass nouns (*This is sugar*). That could explain the results with the adjective frame.

but *He kisses dog*, omitting the determiner because the weak syllable is taken up by the plural *–(s)es*. This phonological approach is not incompatible with the incomplete knowledge approach sketched above, and might well complete it.

4.4. Word classes

Children must not only determine whether a word is a noun or not, but also, if it is a noun, which type (subcategory) of noun it is. English distinguishes between count nouns (*car*), mass nouns (*sand*), and proper nouns (*Peter*). Count nouns occur with determiners like *a, many, several*, mass nouns follow determiners like *some*, *much*, and proper names are not preceded by determiners.

Here again, a number of studies have shown that children are aware of the distinction surprisingly early. In a classic study, Katz, Baker and Macnamara (1974) presented 17-month-old children with a novel word for a doll. For some children, the noun was preceded by a determiner ("This is a wug"), for others, it was not ("This is wug"). The children tended to interpret the word preceded by a determiner as naming the kind of doll, and the bare noun as giving the name of the specific doll. This distinctive behavior was mainly observed with girls, and it occurred only when the object labeled was a doll (i.e. person-like), and not when it was a block. This study taps the beginnings of the capacity of children to take into account the presence or absence of the determiner, and to use it as an indication that a word is a common noun or a proper noun (see also Gelman and Taylor 1984). Soja (1992) got similar results with the distinction between count and mass nouns with two-year-olds. Such experiments show not only that children categorize nouns on the basis of the cue provided by the determiner, but also that they use what they know of the denotation of a mass noun or a

count noun to infer the meaning of words. Thus distributional cues are actively used by children in interpreting sentences (a strategy called 'syntactic bootstrapping' by Gleitman 1990, in the context of the acquisition of verb meanings). But we should guard against attributing too much competence to the children: when form class cues conflict with word learning constraints like the whole-object assumption, children often ignore the formal cues (Woodward and Markman 1998).

While the subcategories of nouns are learned early, other word classes are late acquisitions. In some cases, learning a subcategory of words requires establishing a dependency between words. For example, a consistent crosslinguistic finding is that regular pronouns (*him*) are learned later than anaphors (*himself*). Preschool children know that in *John hurts himself*, *himself* must refer to the subject John and not to someone else; but they think that in *John hurts him*, *him* may also refer to John (see Guasti 2002, Thornton and Wexler 1999, for discussion and references). Here the subclass of pronoun determines its interpretative possibilities with respect to a potential antecedent within the clause. In order to learn the distinction, children must consider dependencies between non adjacent words.

Verbs are also subcategorized into various classes, e.g. transitive, intransitive. One intriguing problem is that of transitivity alternations illustrated in (6), and repeated below.

(12) a. John broke the vase.

b. The vase broke.

The alternation is limited to certain subclasses of verbs. Verbs of manner of motion alternate (*swing, slip*), but verbs of directed motion don't (some are intransitive : *rise, fall* ; others are transitive : *pull, raise*). How do children learn these *narrow conflation classes*? Pinker (1989 :270) argues that traditional category formation would not work because there is no simple definition that would include only the classes of verbs that alternate and exclude those that don't. He proposes to equip the child with innate knowledge of the types of semantic structures available for verbs; upon learning that a verb alternates, the learner would assume that other verbs with the same semantic structure alternate. Braine and Brooks (1995) rather propose that children adhere to a "unique argument-structure preference" principle: once one argument structure is firmly learned for a verb, it tends to preempt other argument structures for that verb until the language they hear teaches them otherwise. This view, which would explain the rarity of transitivity errors (**I'm falling it*), finds support in a study by Brooks and Tomasello (1999), who taught children novel verbs in one construction and looked at whether they were ready to use them in the other frame. Children of 2;5 tend to use verbs in the construction in which they are learned. When they produce transitivity alternations, they don't distinguish between alternating and non alternating verb classes. It is only by six or seven years of age that the narrow classes seem to be acquired.

Culicover (1999) argues that narrow range classes of words are ubiquitous in language. For example only a small subset of adjectives can appear in the impersonal construction in (14), and, of this set, a tiny subset can also appear with a similar meaning in the construction in (15) (Culicover 1999 :47).

(14) It is *likely* that Robin will be elected President next year

probable, possible, certain, clear, sure

(15) Robin is *likely* to be elected President next year.

*probable, *possible, certain, *clear, sure

How do children manage to learn these tiny subclasses? According to Culicover, the learner must be a **conservative attentive learner**, who will recognize regularities, but will generalize only if the number and distribution of cases exceeds certain bounds (Culicover 1999: 29-30). Because of the variety of minor formal categories which must be learned, because distributional mechanisms are necessary to learn them, and because the same mechanisms can construct major formal categories, Maratsos (1998:447) concludes that to hypothesize innate knowledge of major categories is not only unnecessary, it is also theoretically unparsimonious. It could however be claimed that the only way children can converge on the proper subclasses is if they have innate knowledge of the grammatical factors relevant to distinguish them.

4.5. Other cues to grammatical category learning

It has been suggested that children might exploit phonetic or phonological cues to learn parts of speech. In English, nouns are more likely than verbs to be stressed on their initial syllable ($récord_N$ vs $recórd_V$). If children notice this regularity, they could exploit it to categorize new words (Gerken 2001 :156, Kelly 1996). In the case of English, it is not clear that this strategy is of great help, because by the time the child has learned enough words to notice the regularity, he or she is probably able to categorize nouns and verbs on the basis of distributional information alone. But this strategy cannot be discounted. Peters (1997) discusses various other phonological and prosodic factors

influencing the acquisition of grammar, in particular stress, saliency, and rythm (also Morgan 1986, Gleitman and Wanner 1982, Jusczyk 1998, 2001).

5. Models of distributional learning

The evidence reviewed thus far shows that children are sensitive to distributional information. Before the age of two, they attend morphosyntactic elements and use them as cues to determine the grammatical category of words. In this section we review some distributional learning mechanisms that have been proposed to account for the acquisition of grammatical categories.

An early model of a distributional learning mechanism is that of Maratsos and Chalkley (1980, Maratsos 1982). The child registers the semantic, morphological, and distributional properties of lexical items. Once one pattern has been established with one lexical item, other lexical items can assimilate to this pattern, and other patterns can assimilate to this lexical item. As more and more lexical items are learned, the recurring properties of grammatical classes emerge. In Bates and MacWhinney (1987)'s **Competition model**, children attempt to map grammatical forms with their semantic function by taking into account semantic, morphological, and positional cues. In Brent (1996)'s **autonomous bootstrapping** model, children extract a tiny bit of linguistic knowledge in any domain from unanalyzed inputs, and use that bit of knowledge to perform a little more linguistic analysis on future inputs, thereby extracting more knowledge from them (Brent 1996: 25; Cartwright and Brent 1997). As more and more knowledge is extracted, the number of cues that children can use to parse incoming input grows exponentially.

Important notions of cue-based learning models are **cue validity**—the availability and reliability of some cue in the input (an objective notion), **cue strength**—the weight that the organism attaches to some piece of information (a psychological notion), and **cue cost**—the cost of treatment of some cue (Kail 2000). The cues don't need to have 100% validity. Learning depends on discovering new cues and modifying the strength of known ones. The various cues are probably learnt independently of each other, each with their own representational content (Rispoli 1999 :234). They may reinforce each other or, in some cases, compete with each other (Dockrell and McShane 1990).

Neural (connectionist) networks can be viewed as cue-based distributional learning mechanisms computing cue strength as a function of cue validity. They basically perform pattern association. Connectionist models can be quite successful at leaning various aspects of language (Rumelhart & McClelland 1987, Plunkett 1995), but Pinker has repeatedly argued that pattern association is not sufficient to account for grammar learning (Pinker & Mehler 1989, Pinker 1999, Pinker & Ullman 2002). Pinker & Ullman (2002) and Palmer-Brown et al. (2002) also claim that recent connectionist models are really hybrid models of parsing, where some learning is associative, but structure is built in the model (see also Elman et al. 1996 ; Karmiloff-Smith et al. 1998, Karmiloff-Smith 1992). In this context, one interesting hypothesis would be that a connectionist-type distributional learning mechanism classifies new information and categorizes it with similar elements, and that at some point, an internal process of representational redescription (Karmiloff-Smith 1986) extracts the general rule from the knowledge base and stores it in a distinct manner.

6. Constraining the search space

A number of authors have pointed out the difficulty of postulating an unconstrained distributional learning model, because of the enormous number of surface correlations that would need to be computed (Bloom 1999 :287, Pinker 1987, Maratsos 1998:447). Gordon (1988) calculated that the child would have to sift through 8 billion possible contexts to learn the mass/count distinction, yet children command the mass/count syntax by 2;6. What are the constraints on the learning system that allow children to converge rapidly on appropriate linguistic categories? What allows children to home in so quickly on the relevant cues?

Among the proposed solutions to this problem is of course the idea that children would come equipped with **innate linguistic knowledge**, with a 'restricted property register' (Maratsos 1998) considering only certain properties as grammatically relevant. One might also think of **innate processing constraints**. Children might be geared towards looking for some types of contexts and not others. For example, a child would register whether a determiner agrees with the noun, but will not try to compute whether the noun agrees with the last morpheme in the sentence (Pinker 1987). A similar effect is obtained if children restrict their search to phonological phrases (Peters 1997). Aslin, Saffran and Newport (1999 : 378) propose to replace innate constraints by **innately biased statistical learning mechanisms** but they don't specify what these innate biases are. Other authors focus on the **attentional span** of the learner. Redington, Finch and Chater (1998) show that registering adjacency relations is a powerful learning mechanism; Gerken and McIntosh (1993) propose that children focus only on the location of words relative to frequent morphemes (also Valian and Coulson 1988). In the

same vein is Elman (1993) showing that a connectionist model succeeds better if its initial input is limited. Addressing this problem experimentally, Santelmann and Jusczyk (1998) have started to explore the capacity of children to identify dependencies between elements such the auxiliary *is* and the ending *-ing* on the verb (*the dog is running* but not **the dog can running*). They find that 18-month-olds, but not 15-month-olds have developed a sensitivity to this dependency. They also show that the capacity to track the dependency remained when a 2-syllable adverb intervened between the auxiliary and the verb (*John is always running*), but not when the adverb had 3 or 4 syllables. This suggests that children initially register only adjacency-type relations, and, as their attention and short-term memory span develop, they come to be able to notice non-adjacent depencencies and to track them across more and more distant positions.

The correct solution is probably a combination of these factors, and possibly others.

7. Conclusion

To sum up, it is clear that language input orients the child towards discovering the semantic and syntactic categories coded by the language it is learning. Children exploit every possible source of information at their disposal to determine the meaning and the grammatical category of a word. They pay attention to formal cues from the beginning. The initial cues used are probably partial and tentative. Children would initially notice statistical regularities and use them as cues in parsing language, without necessarily having a sufficient mastery of these cues (or confidence in them) to be able to use them in their productive language. Children may begin with noticing some cooccurrence between adjacent elements, or between some linguistic element and some aspect of the

situation, and use this crude cue to help uncover other potential cues. The strength of each cue would vary according to its ability to help in producing successful parses. Gradually, children become more able to track dependencies over longer stretches of discourse.

Children make few category assignment errors. They are conservative in learning the grammar of their language; they tend to use relational words in the constructions in which they have heard them used and are reluctant to generalize. Generalizations seem to appear in a second stage, when the similarities of behavior among members of a class are registered.

Many questions remain to be answered: What type of innate constraints are built into the system? Are some cues more salient than others to children? Just how much knowledge do very young children have of grammatical categories? How is this knowledge built up? Gordon (1988) stated that there are two points of interest in studying the acquisition of some distinction: the point of first awareness and the point of complete control. In fact, many other points of interest can be identified. Take the grammatical category of determiners. We can identify the point of first awareness of determiners, the point where determiners start being used in parsing the clause, the point of the emergence of fillers in production, the point when children systematically supply determiners to introduce noun phrases. This point corresponds to adequate basic use, but Karmiloff-Smith (1979) has shown that correct use is not the end of learning, and that it takes many years before determiners are integrated into a system of relevant markers, and before they are used proficiently in discourse, taking into account the pragmatic context. Language acquisition studies of grammatical categories face many years of intensive work.

labelle.marie@uqam.ca

References

Akhtar, Nameera and Michael Tomasello. 1997. Young children's productivity with word order and verb morphology. *Developmental psychology* 33(6) 952-965.

Aksu-Koc, Ayhan and Dan I. Slobin. 1985. The acquisition of Turkish. In D.I. Slobin (ed.) *The crosslinguistic study of language acquisition, Vol. 1: The data* (839-878). Hillsdale, NJ: Lawrence Erlbaum.

Allen, Joseph and Mark S. Seidenberg. 1999. The emergence of grammaticality in connectionist networks. In Brian MacWhinney (ed.) *The emergence of language*. 115-151. Mahwah, NJ : Lawrence Erlbaum.

Anderson, Stephen R. 1985. Inflectional morphology. In Timothy Shopen (ed.) *Language typology and syntactic description. Vol. 3 : Grammatical categories and the lexicon*. 150-201. Cambridge : Cambridge University Press.

Aslin, Richard N., Jenny R. Saffran and Elissa L. Newport. 1999. Statistical learning in linguistic and nonlinguistic domains. In Brian MacWhinney (ed.) *The emergence of language* (359-380). Mahwah, NJ : Lawrence Erlbaum.

Bassano, Dominique. 2000. Early development of nouns and verbs in French : exploring the interface between lexicon and grammar. *Journal of child language* 27, 521-559.

Bates, Elizabeth and Brian MacWhinney. 1987. Competition, variation, and language learning. In B. MacWhinney (ed.) *Mechanisms of language acquisition*. 157-193. Hillsdale, NJ : Lawrence Erlbaum Associates Inc.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J. & Hartung, J. 1994. Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* 21, 85-124.

Benedict, H. 1979. Early lexical development : comprehension and production. *Journal of Child Language* 6 : 183-200.

Beretta, Alan, Carrie Campbell, Thomas H. Carr, Jie Huang, Lothar M. Schmitt, Kiel Christianson and Yue Cao. 2003. An ER-fMRI investigation of morphological inflection in German reveals that the brain makes a distinction between regular and irregular forms. *Brain and language* 85(1), 67-92.

Borer, Hagit and Bernhard Rohrbacher. 1997. Features and projections: arguments for the full competence hypothesis. Elizabeth Hughes, Mary Hughes, and Annabel Greenhill (eds.) *Proceedings of the 21st Boston University Conference on Language Development* 1, 24-35. Somerville (MA): Cascadilla Press.

Bowerman, Melissa. 1973. *Early syntactic development : A cross linguistic study with special reference to Finnish*. Cambridge, England : Cambridge University Press.

Bowerman, Melissa. 1985. Beyond communicative adequacy: From piecemeal knowledge to an integrated system in the child's acquisition of language. In Keith E. Nelson (ed.) *Children's language. Vol. 5* (369-398). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bowerman, Melissa and Soonja Choi. 2001. Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman and S.C. Levinson (eds.) *Language acquisition and conceptual development*, 475-510. Cambridge (UK): Cambridge University Press.

Boysson-Bardies, Bénédicte de. 1996. *Comment la parole vient aux enfants*. Éditions Odile Jacob. Paris. Published in English under the title *How language comes to children.* 1999. Cambridge, MA: MIT Press.

Braine, M.D.S. 1963. The ontogeny of English phrase structure: the first phase. *Language* 39, 1-13.

Braine, M.D.S. 1971. On two types of models of the internalization of grammars. In D.I. Slobin (ed.) *The ontogenesis of grammar*. 153-186. New York : Academic Press.

Braine, M.D.S. 1976. Children's first word combinations. *Monographs of the Society for research in child development*, 41 (serial no. 164).

Braine, M.D.S. 1987. What is learned in acquiring word classes—A step towrad an acquisition theory. In B. MacWhinney (ed.) *Mechanisms of language acquisition*. 65-87. Hillsdale, NJ : Lawrence Erlbaum Associates Inc.

Braine, M.D.S., Brooks, Patricia J. 1995. Verb Argument Structure and the Problem of Avoiding an Overgeneral Grammar. In Tomasello, Michael, & Merriman, William E. (Eds.) Beyond names for things : young children's acquisition of verbs (353-376). Hillsdale, NJ: Lawrence Erlbaum Associates.

Braine, M.D.S., R.E. Brody, P.J. Brooks, V. Sudhalter, J.A. Ross, L. Catalano and S.M. Fisch 1990. Exploring language acquisition in children with a miniature artificial language. Effects of item and pattern frequency, arbitrary subclasses, and correction. *Journal of memory and language* 29, 591-610.

Brent, Michael R. 1994. Acquisition of subcategorization frames using aggregated evidence from local syntactic cues. *Lingua* 92, 433-470.

Brent, Michael R. 1996. Advances in the computational study of language acquisition. *Cognition* 61, 1-38.

Brooks, P.J., M.D.S. Braine,L. Catalano, R.E. Brody, and V. Sudhalter. 1993. Acquisition of gender-like noun subclasses in an artificial language : The contribution of phonological markers to learning. *Journal of memory and language* 32, 76-95.

Brooks, Patricia J. and Michael Tomasello. 1999. How children constrain their argument structure constructions. *Language* 75(4): 720-738.

Brown, R. 1957. Linguistic determinism and part of speech. *Journal of abnormal and social psychology* 55, 1-5.

Brown, R. 1973. *A first language: the early stages*. Cambridge, MA: Harvard University Press.

Cartwright, Timothy A., Brent, Michael R. 1997. Syntactic categorization in the early language acquisition: formalizing the role of distributional analysis. *Cognition* 63: 121-170.

Caselli, M. C., Casadio, P. & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language* 26, 69-111.

Choi, S. & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language* 22, 497-529.

Clancy, P. 1985. The acquisition of Japanese. In D.I. Slobin (ed.) *The crosslinguistic study of language acquisition. Vol. 1. The data*. 373-524. Hillsdale, NJ : Lawrence Erlbaum.

Clark, Eve V. 2001. Emergent categories in first language acquisition. In Melissa Bowerman and Stephen C. Levinson (eds). *Language acquisition and conceptual development*. 379-405. Cambridge : Cambridge University Press.

Culicover, Peter W. 1999. *Syntactic nuts : Hard cases, syntactic theory and language acquisition*. Oxford : Oxford University Press.

Davidoff, Jules. 2001. Language and perceptual categorization. *Trends in Cognitive Sciences*, 5(9), 382-387.

Demuth, Katherine. 1992. The acquisition of Sesotho. In Dan Isaac Slobin (ed.) *The crosslinguistic study of language acquisition. Vol. 3*. 557-638. Hillsdale (N.J.) : Lawrence Erlbaum.

Dockrell, Julie and John McShane. 1990. Young children's use of phrase structure and inflectional information in form-class assignments of novel nouns and verbs. *First Language* 10, 127-140.

Echols, C.H. and E.L. Newport. 1992. The role of stress and position in determining first words. *Language acquisition* 2 : 189-220.

Elman, J. L. 1993. Learning and development in neural networks : The importance of starting small. *Cognition* 48, 71-99.

Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi and Kim Plunkett. 1996. *Rethinking innateness : A connectionist perspective on development*. Cambridge, MA : MIT Press.

Fernald, A., G.W. McRoberts and D. Swingley. 2001. Infant's developing competence in recognizing and understanding words in fluent speech. In Jürgen Weissenborn and Barbara Höhle. *Approaches to bootstrapping : Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition*. Vol. 1. 97-123. Amsterdam/Philadelphia : John Benjamins.

Fisher, Cynthia. 2002. The role of abstract syntactic knowledge in language acquisition : A reply to Tomasello (2000). *Cognition* 82(3) 259-278.

Fortescue and Lennert Olsen. 1992. The acquisition of West Greenlandic. In D.I. Slobin (ed.) *The crosslinguistic study of language acquisition : vol. 3*. 111-219. Hillsdale, NJ : Lawrence Erlbaum.

Gelman, S.A. and M. Taylor. 1984. How two-year-old children interpret proper and common names for unfamiliar objects. *Child development* 55 : 1535-1540.

Gentner, D. 1982. Why nouns are learned before verbs : Linguistic relativity versus natural partitioning. In S. Kuczaj (ed.) *Language development. Vol. 2 : Language, culture and cognition*. (301-334). Hillsdale, NJ : Lawrence Erlbaum.

Gerken, Lou Ann. 2001. Signal to syntax : building a bridge. In Jürgen Weissenborn and Barbara Höhle. *Approaches to bootstrapping : Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. Vol. 1*.147-165. Amsterdam/Philadelphia : John Benjamins.

Gerken, Lou Ann and B.J. McIntosh. 1993. Interplay of function morphemes and prosody in early language. *Developmental Psychology* 29, 448-457.

Gerken, Lou Ann, B. Landau, and R.E. Remez. 1990. Function morphemes in young children's speech perception and production. *Developmental Psychology* 27, 204-216.

Gleitman, L. 1990. The structural sources of verb meanings. *Language acquisition* 1, 3-55.

Gleitman, L. and E. Wanner. 1982. Language acquisition : the state of the art. In Gleitman, L. and E. Wanner. 1982. *Language acquisition : the state of the art*. 3-48. New York : Cambridge University Press.

Golinkoff, R.M., K. Hisch-Pasek & M.A. Schweisguth. 2001. A reappraisal of young children's knowledge of grammatical morphemes. In Jürgen Weissenborn and

Barbara Höhle. *Approaches to bootstrapping : Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. Vol. 1*. 167-188. Amsterdam/Philadelphia : John Benjamins.

Gopnik, A., Choi, S. & Baumberger, T. 1996. Cross-linguistic differences in early semantic and cognitive development. *Cognitive Development* 11, 197-227.

Gordon, P. 1988. Count/mass category acquisition : distributional distinctions in children's speech. *Journal of Child Language* 15, 109-128.

Grimshaw, J. 1981. Form, function, and the language acquisition device. In C.L. Baker & J.J. McCarthy (eds.) *The Logical problem of language acquisition* (165-182). Cambridge, MA: MIT Press.

Guasti, Maria Teresa. 1993/1994. Verb syntax in Italian child grammar: Finite and non-finite verbs. *Language acquisition* 3, 1-40.

Guasti, Maria Teresa. 2002. *Language Acquisition: the growth of grammar*. Cambridge (MA): MIT press.

Hirsch-Pasek, K. and Golinkoff, R.M. 1996. *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E.,& Gutierrez-Clellen, V. 1993. Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language* 20, 523-49.

Jusczyk, Peter W. 1998. Constraining the search for structure in the input. In Antonella Sorace, Caroline Heycock, Richard Shillcock (eds.) *Language Acquisition : knowledge representation and processing*. Special issue of *Lingua* 106 : 197-218. Amsterdam : North-Holland.

Jusczyk, Peter W. 2001. Bootstrapping from the signal. In Jürgen Weissenborn and Barbara Höhle. *Approaches to bootstrapping : Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. Vol. 1*. 3-21. Amsterdam/Philadelphia : John Benjamins.

Kail, Michèle. 2000. Acquisition syntactique et diversité linguistique. Michèle Kail et Michel Fayol (eds.) *L'acquisition du langage. Vol. 2: Le langage en développement: au-delà de 3 ans*. 9-44. Presses Universitaires de France.

Karmiloff-Smith, Annette. 1979. *A functional approach to child language. : a study of determiners and reference*. Cambridge, UK : Cambridge University Press.

Karmiloff-Smith, A. 1986. From metaprocesses to conscious access: evidence from children's metalinguistic and repair data. *Cognition* 23, 2, 95-147.

Karmiloff-Smith, Annette. 1992. *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Karmiloff-Smith, Annette, Kim Plunkett, Mark H. Johnson, Jeff L. Elman and Elizabeth A. Bates. 1998. What does it mean to claim that something is 'innate'? Response to Clark, Harris, Lightfoot and Samuels. *Mind and Language* 13(4), 558-597.

Kelly, M.H. 1996. The role of phonology in grammatical category assignments. In J.L. Morgan and K. Demuth (eds.) *Signal to syntax : bootstrapping from speech to grammar in early acquisition*. 249-262. Mahwah, NJ : Lawrence Erlbaum.

Kinkade, M.D. 1983. Salish evidence against the universality of "noun" and "verb". *Lingua* 60, 25-40.

Macnamara, John. 1982. *Names for things : a study of child language*. Cambridge, MA : Bradford Books/MIT Press.

Maratsos, Michael. 1982. The child's construction of grammatical categories. In Wanner & Gleitman. *Language Acquisition : the state of the art*. 240-266. Cambridge : Cambridge University Press.

Maratsos, Michael. 1998. The acquisition of grammar. In Deanna Kuhn and Robert S. Siegler (eds.) *Handbook of child psychology, vol. 2*, 421-466. New York : John Wiley & Sons.

Maratsos, Michael. 1999. Some aspects of innateness and complexity in language acquisition. In Martyn Barrett (ed.) *The development of language*. 191-228. Hove, UK : Psychology Press.

Maratsos, Michael & M.A. Chalkley. 1980. The internal language of children's syntax : the ontogenesis and representation of syntactic categories. In Keith Nelson (ed.) *Children's language, Vol. 2*. (127-214). New York: Gardner Press.

Mills, D.LL., S.A. Coffey-Corina, and H.J. Neville. 1993. Language acquisition and cerebral specialization in 20-month-old infants. *Journal of cognitive neuroscience* 5 : 317-334.

Mintz, Toben H. 2002. Category induction from distributional cues in an artificial language. *Memory and Cognition* 30(5) 678-686.

Mintz, Toben H., Elissa Newport & Thomas Bever. 2002. The distributional structure of grammatical categories in speech to young children. *Cognitive Science* 26(4). 393-424.

Morgan, J. 1986. *From simple input to complex grammar : Learning, development, and conceptual change*. Cambridge, MA : MIT Press.

Naigles, Letitia R. 2002. Form is easy, meaning is hard: resolving a paradox in early child language. *Cognition* 86(2) 157-199.

Naigles, Letitia R. 2003. Paradox lost? No, paradox found! Reply to Tomasello and Akhtal (2003), *Cognition*, 88(2). In press.

Naigles, Letitia R. and E. Hoff-Ginsberg, 1995. Input to verb learning: Evidence for the plausibility of syntactic bootstrapping. *Developmental Psychology* 31, 827-837.

Nelson, K. 1995. The dual category problem in the acquisition of action words. In M. Tomasello and B. Merriman (eds.) *Beyond names for things : Young children's acquisition of verb meaning*. 57-76. Hillsdale, NJ : Lawrence Erlbaum.

Palmer-Brown, Dominic Jonathan A. Tepper and Heather M. Powell. 2002. Connectionist natural language parsing. *Trends in cognitive science* 6(10), 437-442.

Peters, Ann M. 1983. *The units of language acquisition*. Cambridge (UK) : Cambridge University Press.

Peters, Ann M. 1995. Strategies in the acquisition of syntax. In Paul Fletcher and Brian MacWhinney (eds.) *The handbook of child language*. 462-482. Oxford (UK): Blackwell.

Peters, Ann M. 1997. Language typology, prosody, and the acquisition of grammatical morphemes. In Dan Isaac Slobin, *The cross-linguistic study of language acquisition. Vol. 5: Expanding the contexts*, 135-199. Mahwah, N.J.: Lawrence Erlbaum Ass.

Phillips, Colin. 1995. Syntax at age two : Cross-linguistic differences. In Carson T. Schütze, Jennifer B. Ganger and Kevin Broihier (eds.) *MIT working papers in linguistics* 26 : *Papers on language processing and acquisition.* 325-382. Cambridge, MA : MIT Department of linguistics.

Pierce, Amy. 1992. *Language acquisition and syntactic theory: A comparative analysis of French and English child grammar*. Dordrecht: Kluwer.

Pinker, Steven. 1984. *Language learnability and language development*. Cambridge, MA : Harvard University Press.

Pinker, Steven. 1987. The bootstrapping problem in language acquisition. In Brian MacWhinney (ed.) *Mechanisms of language acquisition*. 399-441. Hillsdale, NJ : Lawrence Erlbaum Associates Inc.

Pinker, Steven. 1989. *Learnability and cognition : the acquisition of argument structure*. Cambridge, MA : MIT Press.

Pinker, Steven. 1999. *Words and Rules: The Ingredients of Language*. New York: Basic Books. London: Weidenfeld & Nicolson.

Pinker, Steven and Jacques Mehler. 1989. *Connections and symbols*. Cambridge, MA : MIT Press.

Pinker, S. & Ullman, M. 2002. The past and future of the past tense. *Trends in Cognitive Science*, 6, 456-463.

Plunkett, K. 1995. Connectionist approaches to language acquisition. In P. Fletcher & B. MacWhinney. *The Handbook of child language*. Chap. 2 (p. 36-72). Basil Blackwell.

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

Radford, Andrew. 1995. Phrase structure and functional categories. In Paul Fletcher and Brian MacWhinney (eds.) *The Handbook of child language*. 483-507. Oxford, UK : Basil Blackwell.

Radford, Andrew. 1996. Towards a structure-building model of acquisition. In Harald Clahsen (ed.) *Generative perspectives on language acquisition*. 43-89. Amsterdam/Philadelphia : John Benjamins.

Redington, M., N. Chater, S. Finch. 1998. Distributional information: a powerful cue for acquiring syntactic categories. Cognitive Science 22(4): 425-469.

Rumelhart, D.E. & J.L. McClelland. 1987. Learning the past tense of English verbs: Implicit rules or parallel distributed processing. In MacWhinney, B. (ed.) *Mechanisms of language acquisition*. Chap. 7 (p. 195-248). Hillsdale, NJ : Lawrence Erlbaum.

Saffran, J.R., R.N. Aslin, and E.L. Newport. 1996. Statistical learning by 8-month-olds. *Science* 276, 1926-1928.

Santelmann, L. and Jusczyk, P.W. 1998. Sensitivity to Discontinuous Dependencies in Language Learners: Evidence for Limitations in Processing Space. *Cognition* 69, 105-134.

Schachter, Paul. 1985. Part-of-speech systems. In Timothy Shopen (ed.) *Language typology and syntactic description. Vol. 1 : Clause structure*, 3-61. Cambridge : Cambridge University Press.

Schlesinger, I. 1971. Production of utterances and language acquisition. In D. Slobin (ed.) *The ontogenesis of grammar*. 63-101. New York : Academic Press.

Shady, Michele E. 1996. Infants' sensitivity to function morphemes. Unpublished doctoral dissertation, State University of New York at Buffalo.

Shipley, E., C. Smith, and L. Gleitman. 1969. A study in the acquisition of language : Free responses to commands. *Language* 45 : 322-342.

Slobin, Dan Isaac. 1997. The origins of grammaticizable notions: Beyond the individual mind. In D.I. Slobin (ed.) *The crosslinguistic study of language acquisition : vol. 5. Expanding the contexts*. 265-323. Hillsdale, NJ : Lawrence Erlbaum.

Slobin, Dan Isaac. 2001.Form-function relations : how do children find out what they are? In In Melissa Bowerman and Stephen C. Levinson (eds). *Language acquisition and conceptual development*. 406-449. Cambridge : Cambridge University Press.

Smoczynska, M. 1985. The acquisition of Polish. In D.I. Slobin (ed.) *The crosslinguistic study of language acquisition : vol. 1. The data*. 585-686. Hillsdale, NJ : Lawrence Erlbaum.

Soja, N.N. 1992. Inferences about the meanings of nouns : The relationship between perception and syntax. *Cognitive Development* 7, 29-45.

Suzman, Susan M. 1996. Acquisition of noun class systems in related Bantu languages. In Carolyn E. Johnson and John H.V. Gilbert. *Children's language, vol. 9* (81-104). Mahwah, NJ: Lawrence Erlbaum.

Thornton, Rosalind and Kenneth Wexler. 1999. *Principle B, VP ellipsis and interpretation*. Cambridge (MA): MIT Press.

Tomasello, M. 1992. *First verbs : A case study of early grammatical development*. Cambridge, UK : Cambridge University Press.

Tomasello, M., N. Akhtar, K. Dodson, L. Rekau. 1997. Differential productivity in young children's use of nouns and verbs. *Journal of Child Language* 24, 373-387.

Torrens, V. 1995. The acquisition of inflection in Spanish and Catalan. *MIT Working Papers in Linguistics* 26, 451-472.

Valian, Virginia. 1986. Syntactic categories in the speech of young children. *Developmental psychology* 22, 562-579.

Valian, Virginia and Seana Coulson. 1988. Anchor points in language learning : the role of marker frequency. *Journal of memory and language* 27, 71-86.

Waxman, Sandra R. and Amy E.Booth. 2001. Seeing Pink Elephants : Fourteen-Month-Olds' Interpretations of Novel Nouns and Adjectives. *Cognitive Psychology* 43, 217-242.

Waxman, Sandra R. and D.B. Markow. 1995. Words as invitations to form categories : Evidence from 12- to 13-month-old infants. *Cognitive Psychology* 29(3), 257-302.

Woodward, Amanda L. and Ellen M. Markman. 1998. Early word learning. In Deanna Kuhn and Robert S. Siegler (eds.) *Handbook of child psychology, vol. 2*, 371-420. New York : John Wiley & Sons.

Table 1—Parts of speech and word classes

Parts of speech	Word classes	Examples
N	masculine	le soleil ‘the sun’
	feminine	la lune ‘the moon’
V	auxiliary	have, will
	lexical	stative : *John is resembling Peter
		active : John is jumping around
A	attributive	the ball is red
	non attributive	*the president is former
Adv	manner	he runs quickly
	degree	*he runs very
D	definite	the ball
	indefinite	a ball
P	contentful	around
	grammatical	of

Table 2—Inflectional classes. German masculine nouns.

	Der Tag (the day)	Der Bär (the bear)	Der Vogel (the bird)	Der Uhu (the uhu)
Nominative Plural	Die Tage	Die Bären	Die Vögel	Die Uhus

Table 3—Some two-word utterances and their possible analyses

Child utterance	Part of speech description	Semantic characterization	Pivot grammar
All fix	Adv+V	attribute+object	Pivot+Open
See baby	V+N	action+object	Pivot+Open
More cookie	Adv+N	recurrence+object	Pivot+Open
Boot off	N+Part.	object+action	Open+Pivot
Mommy shoe	N+N	possessor+possession	Open+Open

Table 4—Syntax-semantics correspondences according to Pinker 1984

Grammatical element	Semantic inductive basis
Noun	Name of person or thing
Verb	Action or change of state
Adjective	Attribute
Preposition	Spatial relation, path, or direction

Table 5—Typical markers accompanying lexical categories (based on Slobin 1997, 2001)

Category	Typical markers
Verb	person, number, tense, aspect, mood, potential, desiderative, causation, voice, subject and object agreement, speech act type, social status of interlocutors, speaker's evidence for making claim, ...
Noun	number, gender or word class, case, definiteness, animacy, form, orientation, location with respect to speaker, ...

Subject index:

language acquisition, part of speech, word class, grammatical category, inflection, subcategorization, distributional learning, cue-based learning, connectionism, child language, innateness