Evidence Children Use: Learnability and the Acquisition of Grammatical Morphemes

Author(s): Lise Menn

Proceedings of the Twenty-Second Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Role of Learnability in Grammatical Theory (1996), pp. 481-497

EVIDENCE CHILDREN USE: LEARNABILITY AND THE ACQUISITION OF GRAMMATICAL MORPHEMES*

LISE MENN
University of Colorado

The learnability idea was a great insight: Whatever Language is, it must be something that people can learn, and whatever people are, they must be the kind of organism that can learn language. But there's a problem with using this insight: we don't know enough about either what Language is or what Learning is to make dramatic leaps based on such reasoning. At present, the state of the art is more like having two climbers roped together, both holding on with rather inadequate equipment. Better to inch up in parallel, and keep testing how well the theories on both sides are anchored in data.

Human language learnability, first of all, is not a purely mathematical issue. A model, mathematical or otherwise, has consequences for the real world ONLY to the extent that it is a 'good enough' model of the world; that is, only to the extent that its assumptions match (empirical) facts. (For example, Euclidean plane geometry matches the shape of the world well enough for local navigation on our spheroidal planet, but it won't describe the air route to Japan.) A basic type of learnability argument runs like this: Assume that Language is a system of type T, and that the learner hears it under conditions C. If learners don't have property P—for example, access to an innate grammar—then they will make errors from which they cannnot recover. Therefore, P must be a property of the learner.

Yes—IF the assumptions match reality: that is, if language is indeed of type T AND the conditions really are C, AND the learner doesn't have some other property P′ that will enable it to recover from these errors. Whether the assumptions of an argument are a close enough match to reality, however, is always an empirical matter. Morgan (1989:352) has a useful statement: 'Learnability results utilizing Gold's (1967) framework incorporate assumptions about three components: the amount of information available in language input, the extent of constraints on grammatical hypotheses that the learner may entertain (i.e. the richness of Universal Grammar),[1] and the nature of the psychological mechanisms by which a specific grammar is acquired upon exposure to input.'

I argue in this paper that the assumptions of well known learnability arguments do not give a 'good enough' fit to the real world to help us in understanding either Language or its acquisition, because they do not take proper account of a great deal of the evidence actually available to and used by children (i.e. the conditions are in fact not C), and because they also ignore indications that the learner indeed has other properties which permit recovery from error, among which are psychological mechanisms sensitive to frequency item occurrence. (Doug Roland has further pointed out to me that the implications of the Gold mathematical results themselves are typically overstated,[2] but that's a matter for another paper.)

On the other hand, I must also argue that most studies claiming to show that there is enough evidence in the input for children to learn language 'bottom up', that is, without drawing on an innate grammar, such as those of Moerk (e.g. 1991) or of Bohannon, MacWhinney & Snow (1989), seriously minimize the ambiguity of the information available to the child. Although I disagree with some of his views, Marcus (1993) is correct about that. Confronting that ambiguity and its implications will, I claim, actually strengthen the bottom-up 'language can be learned without an innate grammar' arguments.

I think the data that I present on how two children recover from certain of their errors support the following conclusions:

(a) They recover from these errors on the basis of evidence available in the language rather than because of the maturation of some aspect of the grammar;

(b) In order to use this evidence, they must possess a powerful but fairly general instance-based abstracting capacity, a Language-Making Capacity, of the general type envisioned by Slobin 1985, Bybee 1995, or Bates & MacWhinney 1987 (see also Bowerman, this volume).

Before going on, we should consider use of the terms 'negative' vs. 'positive' evidence in learnability theory, because this terminology creates damaging misunderstandings. 'Negative (syntactic) evidence', in the strict mathematical sense, is evidence that a particular string is ungrammatical, i.e. syntactically illegal, not allowed in the language. But information that leads to the same conclusion gradually—information ends up cumulatively telling the child that a form she has been using is not part of the language—may not contain any single element that can in itself be called 'negative'. This is because 'positive' evidence, which means evidence as to what people DO say, i.e. examples of syntactically legal strings (paired with their meanings), is capable of giving rise to negative feedback, feedback as to what people DON'T say (to express those meanings). For example, continued parental use of *I did it* in contexts where the child would have said *Me did it* gradually makes it clear that *Me did it* is strongly dispreferred. Chomsky (1981:9, quoted in Lightfoot 1989) ackowledges this by referring to the accumulation of positive evidence as 'a kind of negative evidence'. Such cumulated positive evidence is generally called 'indirect negative evidence', and no one disputes its availability to the child.

Marcus (1993:55, fn.) points out that indirect negative evidence 'depends on reanalysis of positive evidence based on mechanisms INTERNAL to the child, rather than input external to the child' (emphasis added), and so in this way it is like positive evidence. Further confusion arises from our common-sense feeling that the overt corrections and revisions like *No, say 'I did it'*—however rare they may be—are negative in the ordinary language sense of that word—that is, they must tell the child that something about what she said was WRONG. Certainly, they do. But they do not, in themselves, tell the child that what she said was UNGRAMMATICAL. Therefore, I agree with Marcus that such corrections are still more akin to 'indirect negative evidence' than they are to negative evidence in the mathematical sense. However, to reiterate the plan of this paper, I claim that two aspects of indirect negative evidence are systematically neglected: those who wish to argue for the need for innate grammar do not sufficiently credit the power of indirect negative evidence, and those who argue for a bottom-up low-innateness approach to language acquisition generally fail to consider the seriousness of the challenges that children face in utilizing indirect negative evidence. I will try to clarify both of these issues, and I will argue, along with such theorists as MacWhinney and Bates (e.g. Bates & MacWhinney 1987), that indirect negative evidence is a central factor in enabling the child to learn what is not in the language.

I should also clarify the sense in which I use the phrase 'to learn language'. I mean 'to induce the grammar, phonology, lexicon, and pragmatics of a language from data available in the environment via unconscious processing and storage mechanisms.' These mechanisms are unavailable to introspection and bear little or no resemblance to metalinguistic abilities, such as explicit verbatim memorization. I assume that these processing and storage mechanisms may be quite strongly biased to extract certain kinds of structures, but this bias does not constitute access to some

universal innate (therefore unlearned) grammar which contains such structures. One source of such biases is surely the human infant's conceptual architecture—for whether one is more concerned with language as a vehicle for communication or for representation, Braine's (1992) position that language is a rough reflection of 'the language of thought' (or, I would say, of some portion of it having to do with actual and potential states of affairs) seems reasonable, as does his contention that this conceptual architecture (to which I would add a number of affective and social categories) provides a basis for initial categorizations of words and structures in the ambient language. Having such a semantic first approximation to the kinds of things people mean with language reduces the labor of the Language Making Capacity, and should immensely increase the rate of its convergence on a 'good enough' grammar.

Most of the data for this paper were published in Peters & Menn 1993: longitudinal studies of Seth and Daniel, two-year-old English-acquiring boys, focusing on the emergence of morphosyntax; the remainder are from unpublished materials on the same two children, who showed very different patterns of the emergence of English 'grammatical morphemes'.

I will first examine some rules that must have been learned from evidence because they are specific to English. Then I will examine data from conversations with both children which illustrate the ambiguity of any particular instance of the feedback information available to them: nothing tells them explicitly when it's their grammar which is at fault, let alone how to repair it. I then will argue that it takes impressive data recording and statistical processing abilities to extract rules from data like these. Therefore, we can assume that those learning abilities are present in all normal children. By Occam's razor, unless we have evidence to the contrary, we should further assume that everything that COULD be learned from evidence by using those abilities IS in fact learned in that way. In summary, we should not claim that an article of knowledge—such as a language's preferred word order—is innate or dependent on setting innate grammatical parameters until we have demonstrated that it CANNOT be learned by these powerful abilities, which must be present just to handle the sort of language particular information that everyone agrees has to be learned.

What evidence do we have from children's behavior that they use evidence from the ambient language? At the level of the instantiation of words, more generally morphemes, there is no disagreement: morphemes—or at least morphs!—are learned, because these are language particular. However, what keeps children from assuming that, say, the *foots* that they say and the *feet* that they hear are equally good options? It can't be utterly impossible for this to be the case—consider American English alternate past tenses *dived* and *dove*—but it is extremely rare. Some researchers hold that innate uniqueness constraints on morphs, such as those suggested by Eve Clark (1987), are necessary; others (e.g. Bates & MacWhinney 1987) hold that the competition inherent in a connectionist learning mechanism will be adequate to give the 'uniqueness' that generally holds. Regardless of the outcome of that debate, my aim here is to review the problem of obligatoriness vs. optionality, because the idea that children can't use evidence to LEARN that something is obligatory motivates many innateness arguments, including the argument for the 'pro-drop' parameter.

The argument for setting a 'pro-drop' parameter goes like this: Children can't simply learn from positive evidence that English and similar languages require overt grammatical subjects (in the overwhelming majority of contexts), because positive evidence can only tell the hearer that subjects are optional, not that they are

obligatory. This is logically correct; but the fact is that people don't behave according to such Aristotelian logic. As philosophers have pointed out for centuries, people typically behave as though they possess categorial, certain knowledge about the world in many areas, from science to politics, even though any evidence that they can have collected cannot justify such behavior within a True/False logic. HUMANS MAKE CATEGORIAL JUDGEMENTS (e.g. right vs. wrong, your theory vs. my theory)—throughout cognition, not just in language—ON THE BASIS OF PROBABILISTIC INFORMATION, even though this is not in accordance with classical logic.

Why do people act in this way? The reason must lie in the properties of the mind. And the difference between an Innate Universal Grammar approach and a Language-Making Capacity approach to explaining how OBSERVED PROBABILITIES about language use become BEHAVIORAL CERTAINTIES about grammar lies in what KINDS of 'mental properties' are postulated to explain the subjective certainty that we experience in making (a large class of) grammaticality judgements.

My principal thesis in this paper, then, is that these 'mental properties' include powerful statistical and structure building information processing capacities (which are probably general cognitive properties). Such capacities must be ascribed to the Language-Making Capacity just in order for children to acquire language particular information. Now since language particular information is necessarily learned on the basis of evidence from the ambient language, if such processing capacities characterize the learning of language particular information, nothing prevents us from supposing that they apply in all aspects of language learning. And if they do, then the need to appeal to innate grammar is clearly reduced, and may well be entirely unnecessary.[3]

If one wants to make an argument that a child is using evidence, merely demonstrating the change of his or her rule toward the corresponding adult rule is not enough—perhaps, it could be argued back, the change is due to maturation, or new access to some aspect of innate knowledge. To have a valid argument that a child has used evidence in making a rule change, one must show that the change in question cannot be due to innate factors alone. Not only must the adult model pattern in question be language particular, but the child's earlier pattern should be similar to patterns found in some known adult language, so that it is compatible with a possible grammar, and so that moving away from the earlier version can't be explained by appealing to the maturation of a constraint. Let's consider Daniel, focusing on his several recoveries from incorrect rules for the use of plural and possessive markers, originally presented by Peters & Menn (1993).

I will show that Daniel revised and re-revised the details of the English plural marking rule. He must have made these revisions on the basis of evidence, and he appeared to be moving from one OBLIGATORY version of the rule to another. He did not maintain 'optionality' as a possibility. Since the change in his grammar was reversed, maturation from a less marked to a more marked version of 'UG' can't be invoked. The only way that such a sequence of events could take place within a strong innate grammar approach would be some sort of evidence dependent reversal of parameter setting, and at that point I think we would have a notational rather than an empirical difference between a Slobin-type Language-Making Capacity theory and a Grammar Maturation theory.

Daniel had had an idiosyncratic early rule from about 2 years to 2 years 3 months: a very 'soft', probabilistic rule, based on the pattern of certain appearances of the several {Z} morphemes in his experience, that word final [s ~ z] is

phonologically conditioned in English (Menn 1971, Menn & Matthei 1992). Then he UN-learned this rule between 2;3 and 2;6. By 2 years 6 months, he showed evidence of productive knowledge of the regular plural and the possessive in English, which were supplied correctly for about a month and a half.[4] Then he began to omit plurals, at first at random: *sleeve down, two prune, two ball, both blue car, blue socks, two blacks, lots sticks, lots boats.*[5]

From 2;6.15 to 2;7.0, Daniel was more systematic, omitting the plurals after modifiers in spontaneous utterances. Table 1 lists all the modifier + plural contexts in the left column, and the naked plural contexts in the right column; note that the pattern appears to be a clean syntactic rule for about a week starting at 2;6.18, although the omissions were always correctable in imitation. This context-sensitive pluralization rule was a somewhat gradual change away from his earlier CORRECT plural rule. The last item for 2;6.23 (*two eyes*) shows the correct rule being reinternalized—or maybe we should say reexternalized. Overgeneralizations of plurals (*mans*) continued during this period.

	AFTER MODIFIER	UNMODIFIED (except *both*)
2;6.0	D: *two blacks* (looking for a second black block)	
2;6.0	*lots sticks out*	
2;6.0	*two blacks?*	
2;6.0	*more/other blacks?*	
2;6.0	*more/other black?*	
2;6.17		*cars in kitchen, mommy?*
2;6.18	D: *two, two, two.*	
	L: *Two what, Danny?*	
	D: [In:o]	
	L: *Two windows.*	
	D: [In:oz]	
2;6.18	*two window*	[wivi] (Stevie) *hold hands*
2;6.18	*more truck*	
2;6.21	*two car, red car* (there were two cars, both red)	
2;6.21		*(l)adies no beard* (don't have...)
2;6.23	(imit.) *four mats*	*both mans*
2;6.23	D: *two (wr)ench*	*buttons*
	L: *Two wrenches.*	
	D: *two (wr)enches*	[uz] *hands, not (f)ork* (formula)
2;6.23		*tea* [gægz] (bags)
2;6.23	D: *two (f)oot*	
	L: *Two feet.*	
	D: *two (f)eet*	
	D: *two (f)eet too?* (of another toy)	
	D: *two eyes* (Note that this dialog did not model the regular allomorph)	
2;6.25		*bought pears?*
2;6.26	*two pin*	
2;6.26	*two* [paps] (tops)	

TABLE 1. Plural markers, Daniel. Adapted from Peters & Menn 1993:759, Table 2. Cf. the adjective marking alternation in the adult model:*two blues /two blue blocks.*

Now consider the possessive, as shown in Table 2. By 2;6, the possessive *'s* had apparently been mastered: as Table 1 indicates, between 2;5.25 and 2;6.10, 4 of 4 required attributive possessive markers were supplied, as well as 2 of 2 predicative possessives. The possessive, like the plural, then underwent a reanalysis between 2;6.18 and 2;7.12. The revision, like that of the plural rule, involved the omission of the {Z} in phrasal contexts:

2;6.24 *Mommy car* vs. 2;6.21 *That Mommy's and Mike's.*

This distribution is the same as the attributive/pronominal distribution of the possessives *my/mine, your/yours, her/hers* etc., being learned during the same period:

2;2.19 *mine cookie?*
2;6.23 *mine bib* (corrected in imitation to *my bib*)
2;6.24 *my chair*

So Daniel also 'learned' that {Z} marks possessive except in attributive N*'s*+N constructions.

The possessive alternation started to break down about three weeks after it first appeared:

2;7.12 [bivz] (Steve's) *pen; Mommy's pen*

	ATTRIBUTIVE		PREDICATIVE/PRONOMINAL	
	NOUNS	PRONOUNS	NOUNS	PRONOUNS
1. Apparent mastery of possessive suffix, 2;6.1–10				
	4/4 with *'s*	none	2/2 with *'s*	1/1 mine
e.g.	*Danny's pa(cifi)er*		*Cake—Jean's.*	*Mine.*
(2;5.15–30, four instances of attributive *mine*; e.g. *mine egg all gone*)				
2. Absence of possessive marker in attributive use; beginning of acquisition of *my/mine,* 2;6.18–23				
	4/4 without *'s*	2/2 mine	9/9 with *'s*	3/3 mine
e.g.	*Daddy pen*	*mine bib*	*Daddy's?*	*Mine?*
(*Mine bib* correctable on imitation to *my bib,* 2;6.23)				
3. Absence of possessive marker in attributive use; acquisition of *my/mine,your/yours,* 2;6.24–2;7.9				
	5/5 without *'s*	5/5 my	4/4 with *s*	2/2 mine
	1/1 your		2/2 yours	
e.g.	*Wevie (c)er(eal) ...*	*My egg ...*	*Weve's*	*(L)eave mine on?*

TABLE 2. Possessive markers, Daniel. Adapted from Peters & Menn 1993:761, Table 3.

However, it persisted variably for several months, much longer than the incorrect plural rule did; possibly it had reinforcement from the pattern of the possessive pronouns.

Regardless of how Daniel came up with these patterns, he learned NOT TO mark something which he had previously marked. Both his incorrect pattern and the correct pattern are possible rules of language. And then, of course, he did eventually learn English—that is, he changed his rules again on the basis of what he was hearing around him, instead of regarding his rule and the ambient rule as alternatives.

Now let's turn to Seth, and the evidence that he gives us against a 'pro-drop' parameter. This is an argument of a different sort: Where Daniel looks like he 'considered' optional {Z} morphemes only briefly, Seth—in production—really does look like he progressed from optionality to obligation in his use of subjects and—for appropriate verbs—objects (as detailed in Peters & Menn 1993, and as Peters 1996 is continuing to document). But the change in no way resembles the sort of discrete event that is called to mind by the phrase 'to trigger a parameter setting'.

Seth's development of grammatical morphemes can be characterized as 'Prosody first, sounds next, functions and meaning later'—making use of what were called 'phonological toeholds' by Peters & Menn (746). Specifically (quoting Peters & Menn, 749–50):

'(a) Distributionally, early production of protomorphemic fillers in each specific position begins sporadically, with slots being filled more and more often, first by filler syllables and then by increasingly accurate renditions. The variability cannot be accounted for by rules at any linguistic level.

'(b) Each slot is filled by distinctive phonological alternation sets. Therefore, Seth has some phonological information about what goes where, but it is only partial and only applied probabilistically.

'(c) Distributional and phonological development interact. We find different developmental histories for subject vs. object pronouns, and for prepositions as opposed to particles.'

Seth seemed simultaneously to have developed two co-occurring preverbal slots: Subject, plus another slot for some sort of 'helping verb'. This latter slot Peters & Menn called 'Protomodal'; it contains catenatives such as *wanna* and *gonna,* modals (*can*), and auxiliaries (*did, do, are, shall*). The fillers for the Subject slot are primarily vowels, while those for the Protomodal slot are nasals. Examples:

1;10.0	/n fwo ə kəp/	'N throw a cup'
	/ə̃ ge ə kəp/ ?	'N get ə cup?'
1;10.15	/m pIk ə fawis/?	'N pick ə flowers?'
	/n si ə bak/ ?	'N see ə bark?' (of tree)
1;11.0	/ə̃ tak/?	'N talk?'
	/wə̃ tak/?	'wan' talk?'
	/m braš ə tif/?	'N brush ə teeth?'
	/ə braš ə tif/?	'ə brush ə teeth?'
	/n tUk ə bæf/?	'N take ə bath?'
	/n šiši ĩ ə padi/?	'N shishi in ə potty?'

The developmental profiles for the pair of preverbal slots are based on an analysis of 150 of Seth's utterances in each transcript over the period from 1;10 to 2;3. For all Seth's sentences in which overt subjects were grammatically required, Peters calculated the percents of productions of preverbal nulls, fillers, recognizable protomodals, pronoun-modal amalgams, pronouns, and full NPs. These percentages are graphed in Figure 1 (Peters & Menn 1993:752, Figure 2).

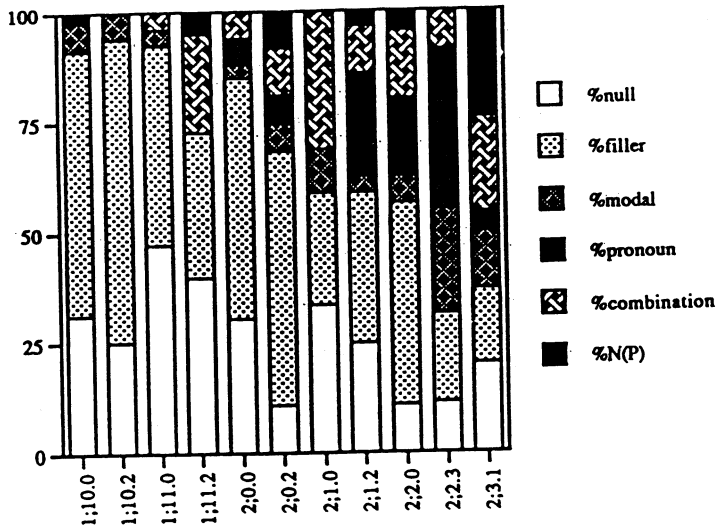

FIGURE 1. Development of preverbal slots: Seth.

Recognizable subject pronouns were infrequent until 2;0.2, and there were very few recognizable subject nouns until 2;3.1, when the noun *Daddy* suddenly occupied nearly 25% of the subject slots, being used in preference to *you* (e.g. *Let Daddy throw it, Wan' Daddy ta bat it*). Attempts to fill both preverbal slots simultaneously were found as early as 1;11.0, but they were rare and at first consisted only of a vowel followed by a nasal (e.g. *ən take ə off*). At 2;1.2 Seth produced a few disyllabic preverbal items such as *nə gə take ə off* (= 'I? go(nna) take it? off') or *əw ə go get it.* Finally, at 2;3.1 Seth seemed to be able to make use of both pre-verbal slots, producing constructions such as *Let Daddy kiss it* or *I can throw again* (see Peters & Menn, 751–52).

This long drawn out variation has none of the crispness suggested by the term 'parameter setting'. I think it supports a Language-Making Capacity approach; but it also requires us to ascribe some serious power to the LMC, and I suggest that a lot of work is still needed to determine exactly what kind of power this must be. The extensive and impressive years of work of the crosslinguistic comparison group led by Slobin give us a pretty good idea of what children learn about grammatical morphemes, and on what time scale; and the differential sensitivities of the psycholinguistic LMC have been made quite explicit over the last 25 years. But the question of how this operates in the communicative environment has barely been broached. That's probably the main thing that the study of language acquisition should be doing for the next 25 years.

Why is it essential to study the acquisition of morphology and syntax 'in the communicative environment'? Let's consider some bits of conversation which show their parents responding to Daniel and Seth by recasting their utterances in different words. These examples will support the statement I made earlier

concerning the ambiguity of the 'correction' that children get—ambiguity from the child's point of view, which is what is important. This ambiguity is in fact just as problematic for ANY instance in which children hear forms different from they have just used or would have used—that is, not just when they hear corrections and recasts, but also when they hear any input speech that they can at least partially understand. Children certainly do learn from language that is not addressed to them directly ('Little pitchers have big ears'—consider the research on children's roleplaying!, e.g. Andersen 1990)—and from the language that is addressed to them which is not a rephrasing of something they have said.

The issue is this: How does the child know when a recasting of what s/he has said is a CORRECTION, and when it's ANOTHER OPTION? If you say *Here book* and your father recasts it as *Here's the bunny book,* how do you distinguish the obligatory grammatical additions from the optional content additions? Although the critics of the bottom-up school of thought have talked about this for a long time (cf. Grimshaw & Pinker 1989), the issue of identifying what is a correction has been a problem that is gravely underacknowledged by bottom-up proponents (e.g. Moerk, Bohannon)—although some of them, e.g. Bohannon, MacWhinney, and Snow (1990), have touched on it a bit.

A very clear statement of the point of view that I share is in Morgan, Bonamo, and Travis (1995:181): 'Child-internal processes of DISCRIMINATION, INTERPRETATION, and RE-EVALUATION must be involved in utilizing corrective feedback, and none of these is trivially accomplished' (emphasis original). Logically, a child must first ascertain what the difference is—in both form and function—between what s/he has said and what the adult has said. The words used by the adult have to be noticed, parsed, and interpreted semantically and pragmatically, or the input is just going to be a string of sounds. Consider the following exchange at 2;7.9 (Menn, unpublished diary data):

D: *Daddy have book* [describing picture]
L: The Daddy has a book.
D: *Daddy have book.*
L: *The Daddy has a book.*
D: *Daddy have<u>s</u> book.*

Daniel repairs one morpheme at a time: note the evidence this provides for his being able to parse what I have said, and to compare at least part of it with what he has said. This parsing ability is evident throughout his responses to recasts (when they have any audible effect), and it's what also happened in Brown's Adam's notorious *Nobody don't likes me* response (McNeill 1966). (Moerk 1991:232, Table 5, gives many more examples of both successes and failures at picking up on parentally-added grammatical morphemes.)

Now consider the following from Daniel at 2;7.11:

D: *Me get down.*
L: *Say: I'm getting down.*
D: *Me get down.*
L: *I'm getting down.*
D: *I get down.*

How does Daniel know, in any of these cases, whether my changed form is a grammatical correction, a formality correction, a politeness correction, a semantic correction? It is no wonder that the immediate effect does not translate into a change in the child's grammar; we are not seeing a system that is IMPERVIOUS to

correction, but rather a system that is (appropriately) CONSERVATIVE in its response. It waits, as it must, to gather a considerable amount of information about why the adult has made a recast before making a change in its morphosyntax, its lexicon, or some part of its pragmatic system, as may turn out to be appropriate. On 2;7.16 we find the following failure to incorporate my correction of five days earlier, just as the literature leads us to expect:

D: *Me diit.*
L: *I did it.*
D: *I diit.*

But Daniel had a lot of work yet to do in order to determine when *I* is preferred to *me*—certainly a wholesale replacement of *me* by *I* in all environments would have been inconsistent with what he had been hearing his elders say.

Four days later (2;7.20), Daniel's grammar indeed starts to change, but the new form is 'optional'—that is, we cannot tell what his unconscious hypothesis might have been as to when it should be used. All we know is that he now spontaneously produced both forms:

D: *I diit.*
D: *Me diit.*

My attempt to correct this *Me diit* by modeling *I did it* resulted in the following dialog, which shows again how hard it is for the child to determine the purpose of an adult's response:

D: *Me diit.*
L: *I did it.*
D: *You do it too?*
D: *You did it?*
L: (weakly) *I helped.*

More dramatic illustrations of the problems of interpretation of adult responses come from Seth, who is visually impaired, and has the blind child's typical prolongation of problems with working out first and second person pronominal reference (Dunlea 1989). This compounds his difficulties in finding out what his father is attempting to accomplish with his recasts and responses in the following conversations. The following are interchanges between Seth (S) and his father Bob Wilson (B), at age 2;5 (from Peters 1966:6–7):

(a) S: *I-I-I wanna get i' really high, Dad.*
B: *You want Daddy ta help you?*
S: *Wan' Daddy to help you, Dad.*
B: *Say, 'Help me, Dad.'*
S: *Help me, Dad.*

(b) S: *I don' wanna* – [crying] *wanna build dat* [crying]
B: *D'you want Daddy ta help you build-tower?*
S: *Wan' Daddy ta help myou build-tower.*
B: *OK, let's build one more.*

(c) S: *Wan' Daddy ta pick you up?*
B: *Y' want me ta pick you up?*
B: *D'ya wanna git outa your high chair?*

How is Seth to know how these forms from his father relate to each other and to his own forms syntactically, semantically, and pragmatically? Which are corrections of incorrect forms? Which are examples of pragmatically preferred alternatives?

I think these data illustrate several important points. First, recasts probably do not have much of a privileged position in helping children learn, regardless of what parents think. Yes, parental models are important, but as many of us (notably MacWhinney and Bates) have long held, and as Morgan et al. (1995:194) observed in their statistical study of recasts in the Brown corpus, 'errors could ... also be revealed by covert comparisons between parental utterances and how the children themselves would have said the same sentences. On this latter view, any positive exemplar could instigate comparable insight. Note that recasts provide mere fractions of the total input for any particular construction ...' (In some of the Roger Brown corpora, Morgan et al., p. 194, found that recasts provided only about 5% of the article models, and 3% of the WhQ+aux models that were actually present in the data.) The key point is that, quoting again (194): 'no recast provides information about why an error is an error ... Children must first guess the nature of the problem—whether it is syntactic, semantic, morphological, phonological, or pragmatic—and must then guess the generality of the required repair.' The child's LMC bears an enormous burden: figuring out WHY the adult has provided a particular recast or model in a particular situation.

Marcus (1993:77) makes the same point, but goes too far beyond it: 'Recasts and expansions may serve as ideal positive evidence, but they do not tell children what is not in the language.' His second clause is true only in the INSTANCE, not in the AGGREGATE. This is because children are not restricted to making grammar decisions one string at a time; instead, they accumulate evidence, as years of documentation of frequency effects tell us. A good statement is made by Snow & Tomasello (1989:357–58): 'A close look at children acquiring particular structures (e.g. the use of prepositions or the past tense of regular verbs) often reveals days, weeks, or even months during which they hear and struggle to comprehend dozens or even hundreds of relevant adult examples, and during which they struggle, on numerous occasions, interspersing both successes and failures, to produce the correct form themselves.' (For other statements and supporting data, see MacWhinney 1987, passim.) While no individual response that children hear could be interpreted by them as the equivalent of 'string not in language', they can cumulatively compare what they hear to what they would have said—and also, as Ann Peters points out (p.c., 1996), responses that they actually get to the responses they expected to get.

In summary: it must be the case that children can in fact learn what is a matter of grammar, what is a matter of politeness, and what is a matter of additional information. But deciding that 'evidence' about what people say is actually 'negative'—something tantamount to a call of 'ungrammaticality'—takes a while: and this had better be the case! (Cf. Marcus 1993:77: 'Children who changed their grammars every time the parent said something different would radically damage their languages.') The so called demonstrations that children do not use negative evidence even when it is present are better seen as demonstrations of how a conservative system deals with the multiple possible functions that a recast might have.

What processing capacity is needed to determine what items are obligatory IN CONTEXT—where 'context' includes meaning and setting? Statistically based processing must be taking place—that is, recurrent pairings of parsed, interpreted strings with a long list of semantic and pragmatic variables. Crucially, the LMC must be capable of constructing categories and of labeling them—the kind of issue on which Maratsos (e.g. 1982) has especially focused—and it must be responsive to the uses of constructions as well.

As regards children's ability to respond to some aspects of form and meaning in advance of their abilities to produce them, we can invoke the findings of Golinkoff & Hirsh-Pasek (e.g. 1987), who have shown that a significant degree of comprehension of the argument structure carried by basic SVO word order for common action verbs is present in the early one word stage. We can also look to Tomasello & Barton (1994), who have shown that social-pragmatic cues are attended to differentially by two year olds in both noun learning and verb learning;[6] to the findings of Gerken and colleagues (e.g. 1989, 1991), who have shown that identification of grammatical morphemes precedes contrastive use in production, typically by many months; and to Bloom, Hood & Lightbown for the evidence that children who imitate adult models 'imitated only word and structures in the speech that they heard which they appeared to be in the process of learning' (1974:416), so that 'When a child imitated an adult utterance he must have already processed it to the extent that he recognized that some aspect of the utterance was in that gray area ... it was not entirely new to him nor already in his productive competence ... [T]his level of processing ... involved recognition of partial relationships among an aspect of utterance form, an aspect of the situation, and some information about form and content already in cognitive memory' (1974:417).

A nice example showing this 'leading edge' character of children's choices of forms to imitate is the following, where Daniel (2;8.9) took a verification question as an occasion to incorporate the modeled form in his response:

D: *Not fits.*
L: *It doesn't fit?*
D: [nʌs] *fit.*

So we see that the child who is hearing a typical recast of her speech is likely to find some of it within her zone of proximal development: that is, she can probably identify the 'extra' grammatical morphemes as such, map out the most frequent SVO action structure, and spot the content words that the adult is putting in. Those abilities are essential to executing the parsing that is needed for the correlational analysis discussed above.[7] And again, this much power is needed just for language specific learning, which by definition cannot depend on an innate grammar.

Returning, finally, to the issue of the categoriality of children's rule changes: Daniel, retreating from his rule forbidding plural marking inside X+N phrases, must have been able to (subconsciously) decide that *two blocks* was correct, and *two block* incorrect. He had to do this on the basis of hearing adults say *two blocks*. But why didn't he conclude that *two blocks* was simply an alternative form? Why abandon his rule-governed *two block* form? There is no escaping the conclusion that he came to a categorial conclusion on the—necessarily probabilistic—basis of the sample of language that he heard up to the point at 2;7 when he stopped omitting the plural markers on the second elements of NP's. This constitutes categorial retreat on the basis of evidence, which children are not supposed to be able to do, according to proponents of innate grammar.

If we contrast Daniel's categorial behavior with Seth's extremely slow and probabilistic progress towards the more complex matter of sorting out the whole pre-main-verb slot, and if we are willing to interpolate between these extreme cases, then we can form a preliminary notion that may bridge the large gap between them. Daniel, during the period we are studying, was collecting data about the distribution of two fully parsed {Z} affixes whose meanings he already understood (it had demonstrably taken months for him to get to this point: see Menn & Matthei 1992). Seth, by contrast, appeared to have as yet very little idea of the parsing or the

semantic and pragmatic factors governing this 'slot' he was working with—cf. the blend *myou* in line 3 of ex. (a), above. He showed increasingly long strings which more and more clearly consisted of grammatically sequenced pronouns and modals, but the learning seemed to be very formulaic—indeed, almost lexical, with continuing pronoun reversals and little or no evidence for semantic differentiation (Peters & Menn 1993, Peters 1996). He must have been working on many fronts at once: if these children were conscious scientists, we'd say that Daniel was checking out a pair of simple well formed hypotheses, which didn't take many samples to yield significant results by simple t-tests, whereas Seth was on a fishing expedition and had to carry out a gigantic factor analysis to get his results.

Now, that kind of statement is a fantasy. But the LMC must have mechanisms which produce these fantastic results. If it turns out that, in general, categorial rule making occurs when the search space is small, and gradual rule making when the search space is large, then we'll have another handle on the nature of the LMC which we can use to constrain our modeling.

So I will put forward for test the following, which I will call the MENN-PETERS CONJECTURE (formulated in discussion with Ann Peters): Quick 'hard-edged' rule change will appear when the variables involved are few, and the LMC's analysis of the units involved is already adequate; protracted 'soft' probabilistic rule change will appear when there are many variables involved, and the LMC's analysis of the units (e.g. placement of morpheme boundaries and identification of allomorphs) is under way simultaneously.

Research to support, refine, or refute this conjecture requires sampling of development that is temporally fine-grained enough to detect swift rule changes (within one week) like Daniel's plural and possessive rules, and extensive enough (at least 90 minutes) at each sampling point to confirm the probabilistic quality of protracted changes like Seth's development of the pre-verbal slot.

Indeed, this relation between the number of variables being treated simultaneously by the child and the cleanness of the rule change may not be language specific behavior at all—it's a good candidate for a general cognitive property of the sort that would have profound consequences for language, so our conjecture should be tested in other areas of cognitive development as well as for language.

How can I argue that one should be willing to countenance all this data processing power operating below the level of consciousness; doesn't it seem simpler to invoke innate knowledge of grammar, instead? Well, no. Making use of an innate grammar will, I think, require the same kind of computational power that I have just argued for. Consider the power required by the parameter setting procedure proposed by Gibson. Or, if you review the discussion of parameter setting in the commentaries on Lightfoot's 1989 article in *Brain and Behavioral Sciences,* you will see that several of the contributors who work in the Chomskyan paradigm are worried about how a parameter could actually be set; they end up saying that the learner must have a very rich analytical ability in order to be able to do any such thing.

For example, Robin Clark (1989:337): 'a parameterized (selective) theory may still face some problems similar to those faced by inductive theories. Put briefly, given a piece of input data, how does the learner know what parameter to set? The learner must have some idea how the datum is relevant to the problem of arriving at the correct adult state.' And Yosef Grodzinsky (1989:342): 'We do not know what prevents a child from assigning a grammatical analysis to "noise", which may lead to wrong conclusions about the grammar of the ambient language.' And Edward

Stabler (1989:360), who asks how it is possible that 'some grammatical strings that occur in a learner's environment could be made grammatical by adjusting the parameters, and yet the learner does not make the adjustment?' At the least, these arguments support O'Grady's concept of 'general nativism', which he characterizes (1989:354) as 'the view that inborn cognitive structures relevant to linguistic development are not specific to the language faculty.'

So I agree with Morgan (1986:196): 'In tandem with observations ... concerning the child's brief, limited exposure to language and the complexity and the abstractness of the grammatical system that the child acquires, these results argue strongly for the importance of properties of the mind in guiding the child's development of language.' And I propose that the 'properties of the mind' in question include a statistically very powerful Language-Making Capacity which can go on a fishing expedition and bring back a good, sometimes categorial catch.

NOTES

* I am happy to acknowledge the substantial contributions of my colleagues; some for providing arguments, some for insisting that I be less confusing, and some for just getting me started, at last, on facing the some of the issues here: Giulia Bencini, Bill Bright, Tom Landauer, Stephen Menn, William O'Grady, Ann Peters, Doug Roland, Valerie Ross, and whoever the folks were at Berkeley that put me on the BLS 22 Learnability panel.

[1] I will not use the term 'Universal Grammar' in my discussion, because it suggests too strongly that what is common to all human languages has the form of a grammar with some details unspecified.

[2] The Gold framework itself is over-restrictive and should be discarded: it crucially assumes arrival at a single correct grammar, rather than an asymptotic approach to any member of set of a number of undetectably different grammars (Feldman 1972). Doug Roland (p.c.) summarizes Feldman's mathematical learnability results this way: 'Using positive evidence only (sans statistics), you can include all of the sentences in the language, reject all incorrect grammars, and guess at the correct grammar for an infinite number of times, if you allow for intervening wrong guesses.'

[3] Giulia Bencini (p.c.) points out to me that, in the special case where the apparent optionality is between presence and absence of a form, rather than between two competing forms, one can apply Paul Smolensky's 'principle of maximal missing information' (1986, based in turn on Claude Shannon's classical work on information theory): 'In guessing the probability distributions of certain patterns within the environment, you assume that the environment has no more dishomogenity than what is needed to account for the given information.' In other words, a system adapted to robust learning through noise should be biased to assume that forms which are perceived as being present sporadically and randomly are in fact 'there' even when not heard. Behavior supporting this as a learning principle is already documented for some children (cf. Thieman 1975, Menn & Matthei 1992), and should be searched for in other studies.

[4] Examples are *two prunes*; but *books, lots boats, cars in, crash cars, too many trucks, slippers off, no seeds, hi guys, okay guys, get cars, no cars, ladies, boys, blue socks.* For plural, overgeneralization indicated full productivity at 2;5.25: [Irz, Irzəz] 'scissors'.

[5] The adjectives marked with plural *s* are used appropriately; they referred to plural colored blocks, sticks or cars.

[6] '...[C]hildren learned a novel noun for an object the adult was searching for, not ones she had rejected while searching', and they learned modeled verbs for actions that appeared to be intentional, rather than accidental (i.e. followed by *Oops!*)

[7] I may have been a bit pessimistic about the cues available in the input language. There may be prosodic cues that a child can learn to use in interpreting what an adult means by recasting his/her utterance; some errors that impede understanding by the adult appear to be clearly marked. Menn & Boyce (1982) showed that verification questions in parental speech to children (in Berko-Gleason's laboratory playroom) were marked by extremely high pitch, as well as having clear corrective content.

Father: *And what's on the door?* [referring to picture]
N: xxx *Sesame Street.*
F: *What's from Sesame Street?*
(falsetto) *Is that Sesame Street?* N: *Yeah.*
F: *Is it?* N: *No.*
F: *Yeh, I didn't think it was.* N. *I think it was.*
F: *Yes, but it's not.*
That's not Sesame Street.

The existence of this semantic type of correction is generally acknowledged, so these examples are only mildly interesting in themselves. The question is: Are other types of corrections also suprasegmentally marked? Contour, tempo, and suprasegmentals in general are grossly undertranscribed, and CHILDES' coding format (CHAT) is unfortunately not well adapted for suprasegmentals, so this is an area that needs technical development in order to be properly studied. To determine whether adults are giving children stress or pitch cues about the nature of the difference between the way the child has said something and the way the adult recasts it, suprasegmental information—at least impressionistic, preferably instrumental—must be available; it should be easy to enter and searchably represented in transcripts.

REFERENCES

Andersen, Elaine. 1990. Speaking with style: The sociolinguistic skills of children. London: Routledge.
Bates, Elizabeth, and Brian MacWhinney. 1987. Competition, variation, and language learning. In MacWhinney 1987, 839–78.
Bloom, Lois; Lois Hood; and Patsy Lightbown. 1974. Imitation in language development: If, when, and why. Cognitive Psychology 6.380–420.
Bohannon, John Neil; Brian MacWhinney; and Catherine Snow. 1989. No negative evidence revisited: Beyond learnabililty or who has to prove what to whom. Developmental Psychology 26.221–26.
Bowerman, Melissa. 1996. Argument structure and learnability: Is a solution in sight? BLS 22.
Braine, Martin D.S. 1992. What sort of innate structure is needed to 'bootstrap' into syntax? Cognition 45.77–100.
Bybee, Joan. 1995. Regular morphology and the lexicon. Language and Cognitive Processes 10.425–55.
Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.
Clark, Eve. 1987. The principle of contrast: A constraint on language acquisition. In MacWhinney 1987, 1–33.

Clark, Robin. 1989. Causality and parameter setting. Behavioral and Brain Sciences 12.321–75.
Dunlea, Anne. 1989. Vision and the emergence of meaning. Cambridge: Cambridge University Press.
Feldman, Jerome A. 1972. Some decidability results on grammatical inference and complexity. Information and Control 20.244–62.
Gerken, Lou Ann. 1991. The metrical basis for children's subjectless sentences. Journal of Memory and Language 30.431–51.
_____ ; Barbara Landau; and Robert Remez. 1989. Function morphemes in young children's speech perception and production. Developmental Psychology 26.204–16.
Gibson, Ted. 1996. Triggering data and parameter setting. Presentation at BLS 22.
Gold, E. M. 1967. Language identification in the limit. Information and Control 10.447–74.
Golinkoff, R.; K. Hirsh-Pasek; K. Cauley; and L. Gordon. 1987. The eyes have it: Lexical and syntactic comprehension in a new paradigm. Journal of Child Language 14, 23–45.
Grimshaw, Jane, and Pinker, Steven. 1989. Positive and negative evidence in language acquisition. Behavioral and Brain Science 12.341–42.
Grodzinsky, Yosef. 1989. The language learner: A trigger-happy kid? Behavioral and Brain Sciences 12.342–43.
Lightfoot, David. 1989. The child's trigger experience: Degree-0 learnability. Behavioral and Brain Sciences 12.321–34.
McNeill, David. 1966. Developmental psycholinguistics. The genesis of language: A psycholinguistic approach, ed. by Frank Smith and George A. Miller, 15–84. Cambridge MA: MIT Press.
MacWhinney, Brian (ed.) 1987. Mechanisms of language acquisition. Hillsdale, NJ: Erlbaum.
Maratsos, Michael. 1982. The child's construction of grammatical categories. Language acquisition: The state of the art, ed. by Eric Wanner and Lila Gleitman, 240–66. Cambridge: Cambridge University Press.
Marcus, Gary F. 1993. Negative evidence in language acquisition. Cognition 46.53–85.
Menn, Lise. 1971. Phonotactic rules in beginning speech. Lingua 26.225–51.
_____, and Suzanne Boyce. 1982. Fundamental frequency and discourse structure. Language and Speech 25.341–83.
_____, and Ed Matthei. 1992. The 'two-lexicon' model of child phonology: Looking back, looking ahead. Phonological development: Models, research, implications, ed. by Charles A. Ferguson, Lise Menn, and Carol Stoel-Gammon, 211–47. Timonium, MD: York Press.
Moerk, Ernst L. 1991. Positive evidence for negative evidence. First Language 11.219–51.
Morgan, James L. 1989. Learnability considerations and the nature of trigger experiences in language acquisition. Behavioral and Brain Sciences 12.352–53.
Morgan, James L.; Katherine M. Bonamo; and Lisa L. Travis. 1995. Negative evidence on negative evidence. Developmental Psychology 31.180–97.
O'Grady, William. 1989. Two perspectives on learnability. Behavioral and Brain Sciences 12.354–55.
Peters, Ann M. 1996. The emergence of catenatives from filler syllables. Linguistics Department, University of Hawai'i, MS.

———, and Lise Menn (1993). False starts and filler syllables: Ways to learn grammatical morphemes. Language 69.742–77.
Slobin, Dan Isaac. 1985. Crosslinguistic evidence for the language-making capacity. The crosslinguistic study of language acquisition, ed. by D. I. Slobin, 2:1157–249. Hillsdale, NJ: Erlbaum.
Smolensky, Paul. 1986. Information processing in dynamical systems: Foundations of harmony theory. Parallel distributed processing, I: Foundations, ed. by David E. Rumelhart and James L. McClelland, 195–281. Cambridge, MA: MIT Press .
Snow, Catherine, and Michael Tomasello. 1989. Data on language input: Incomprehensible omission indeed! Behavioral and Brain Sciences 12.357–58.
Stabler, Edward P., Jr. 1989. What's a trigger? Behavioral and Brain Sciences 12.358–60.
Thieman, Thomas J. 1975. Imitation and recall of optionally deletable sentences by young children. Journal of Child Language 2.261–69.
Tomasello, Michael, and Michelle Barton. 1990. Learning words in non-ostensive contexts. Developmental Psychology 90.639–50.

*****

The author and publishers of BLS 22 would like to thank the Linguistic Society of America for permission to reprint Figure 2 from "False Starts and Filler Syllables: Ways to Learn Grammatical Morphemes" (LANGUAGE 69: 752) as Figure 1 of this paper.