To appear in J. Schaffer and Y. Levy (Eds.), Language competence across populations: Towards a definition of Specific Language Impairment, Earlbaum, 2002

LENNEBERG'S DREAM: LEARNING, NORMAL LANGUAGE DEVELOPMENT AND SPECIFIC LANGUAGE IMPAIRMENT

Ken Wexler, August 10, 2001

In recent years the field of language acquisition has made remarkable progress. In my view, there is no area of cognitive science that has advanced at a quicker pace. The field is full of reliable and non-obvious generalizations, relations to other fields are understood, a good deal about the relation between normal and impaired development is understood and the relative contributions of learning and development have begun to be sorted out in a coherent manner. In this chapter I would like to sketch out some of these results, attempting to give an over-all view of the central questions and the answers that current research suggests. For the domain of phenomena I will pick one important case -- the development of central properties of sentence structure. I have sacrificed breadth of coverage in phenomena and precision of technical development in order to have space to discuss the central questions and to make the results available to non-specialists. A major purpose of my paper is to show how important discoveries concerning impaired linguistic development (SLI), one of the foci of this book, flow naturally from and contribute to the advances in the study of normal language acquisition.

The Computational System of Language

The area of language with which I will be concerned is a central one; what Chomsky (1995) calls the *computational system* of language. This is the part of language internalized by the mind/brain that is responsible for basic properties of sentence construction -- that part of language that is central to the conveying of complicated and non-context-bound ideas, that part of language that seems bound up with species-specific biological properties. A more traditional term for this part of language is *grammar*. I will concentrate on sentence grammar, mostly ignoring phonology, the lexicon and pragmatics.

The properties of the computational system (grammar) that I will discuss include properties of syntax and semantics (ignoring phonology). There is more known about the development of syntax, and I will concentrate on that area, although we will discuss semantics at different points.

The computational system of sentence grammar has two parts. First, there are *Principles* that hold of all languages; by hypothesis they are computed by the brain as the result of genetically guided mechanisms. Second there are *Parameters* that are set differently by different languages. The parameters are set by an individual as the result of experience. The idea is that they can be set simply from experience, so that a child can easily hone in on the correct grammar given that she has the principles.[1]

This framework poses the following questions for the study of the development of language.

(1) a. How does the computational system of language develop?
b. What is learned?
c. What is genetically guided?
d. What develops late under genetic guidance?
e. What kind of variation in development is there across languages?
f. What kind of impairment occurs in development and learning?
g. What does variation across languages and individuals tells us about the genetic structure of language?

These questions are among the most central questions for any biological system that is influenced both by genetics and by the environment, as language is. It is hard to see how we can make progress on other developmental questions in the cognitive (neuro-)science of language without finding reasonable answers to these questions.

One way to begin to think about the above questions is to ask: what accounts for child errors, for non-adult language? In particular, we would like to know if errors are caused by the mis-setting of parameters or by the growth of mature forms of principles. That is, we would like to know the answers to (2):

(2) a. Are parameters sometimes mis-set by children?

[1]A long series of works shows that the problem of parameter-setting is not so obviously simple as the theory postulates. See Wexler and Hamburger (1973), Wexler and Culicover (1980), Manzini and Wexler (1987), Wexler and Manzini (1987), Clark and Roberts (1993), Gibson and Wexler (1993), Dresher (1999), Fodor (1998) Bertolo et al (1997) and many other works. Nevertheless, as we will see, since children set their parameters so well and quickly, there should be a simple solution.

b. Do some principles take time to develop in their mature adult form?

It is quite natural to suppose that parameters are sometimes mis-set by a child and that this mis-setting leads to observed child-errors (2a). After all, we know that parameter settings are at least partially the result of experience -- different languages have different parameter-settings, and the only possible way that these parameter-settings could be attained is via learning. As far as I know, there is no evidence that variation in normal language development is genetically-linked, that is, that children from a long genetic background of Italian speakers find learning Italian easier and that children from a genetic background of Chinese speakers find learning Chinese easier. The fundamental empirical result is that any normal child can easily learn any natural language[2].

We could thus understand child errors as the result of difficulties in the process of parameter setting. Such a hypothesis has often been made. In fact, many developmental psycholinguists have assumed that errors in setting parameters were the only errors that children made in developing syntax, outside of errors in learning the lexicon. (See Wexler 1998 for discussion).

It is also quite natural to suppose that principles take time to develop in their mature adult form (2b). After all, biological organisms generally develop over time; their mature forms are different from their immature forms. This development is in central cases taken to be genetically-guided, although influenced by the environment, but not so much as to alter the central character of the development. In fact, the problem of *development* has often been taken to be the central problem of biology.

The answer to (2b) can tell us much about the developmental structure of the genetic system of language. We will have to answer (2b) at least partially before we can understand, for example, how genetics is involved in the common observation that children do make errors in language at an early age, that their systems are at to least some extent non-adult.

The structure of very early child sentences (up to about 3;0) tells us a great deal about the answers to the questions in (1) and (2). . A tool that I will constantly use (outside of empirical investigations) is the simple confrontation of different

[2]There has been almost no formal empirical study of this question, and it would not be totally inconceivable that in fact there is some genetic linkage to linguistic variation. Given the common experience that a child brought up in a language easily learns that language, any genetic linkage would be expected to be extremely subtle.

possible answers to these questions, asking whether these answers can or cannot predict the empirical results. Common-sensical as this tool is, it has only occasionally been used in past understanding of linguistic development, where *a priori* hypotheses have often been taken for granted, without consideration of the empirical facts or the alternative possibilities. For example, it has often been assumed without empirical argument that all errors in child language are due to errors in learning. My strategy here is to keep all reasonable general answers open, arguing for one or the other on the basis of confrontation with evidence.

Inflection and Tense

In this section I will describe some very simple properties of simple sentences. One central property of sentences is that they often have *tense*. Tense is the category which encodes certain time relations. For example, in English we have 2 tenses, *present* and *past*:

(3) a. Mary likes candy (*present* tense)
b. Mary liked candy (*past* tense)

In many) languages, tense is indicated by an "inflection" on the verb, for example the *s* on *like* in (3a) and the *ed* on *like* in (3b). Tense is a grammatical category; it is not the same as *time* For example, in (4) the time of leaving is taken to be next week, the future, but the tense in English does not distinguish a future tense. (It does in some languages).Tensed verbs are often called *finite* verbs. Untensed or non-finite verbs also exist, as in *go* in (5).

(4) Mary leaves next week
(5) Mary wants Bill to go

There is much more to the finite/non-finite distinction that the encoding of tense and the form of the verb. Finite and non-finite verbs behave very differently in many languages.

I will give just one example of this central role of tense, an example that will soon be of use in describing children's behavior. Many languages are what is called *verb-second* (*V2*) languages. This means that in simple clauses, or the main clauses of more complex sentences, the verb always appears in second position although it does not appear there in most clauses. For example, in Dutch (Dutch examples are from Wexler, Schaeffer and Bol in press), verbs usually appear at the *ends* of sentences; Dutch is what is called a *verb-final language*. But in main clauses, the finite verb appears in second position.

(6) morgen gaat Saskia een boek kopen
tomorrow goes Saskia a book buy
ADV V_{fin} SUBJ OBJ Vnonfin
'Saskia is going to buy a book tomorrow'

V_{fin} indicates a finite verb; V_{nonfin} indicates a non-finite verb. Non-finite verbs occur at the end of a clause in Dutch. This is why *kopen*/*buy* appears at the end of the clause. On the other hand, *gaat/goes* is a finite verb; it is marked for present tense, and it appears in 2nd position in the sentence. Either the subject *Saskia* (a name) or the object *een book/a book* could have appeared in first position instead of the adverb. But the verb could not have appeared there.

We can mostly tell in Dutch that a verb is finite or non-finite by its inflection or ending. We know that *kopen* is non-finite because the verb is the root *koop* plus the non-finite (or infinitival) ending *en*. *gaat* on the other hand shows the typical *t* ending of the third person singular present tense (the subject *Saskia* is third person singular).

Syntacticians understand that the finite verb of main clauses *moves* (from final position) to second position, but we don't have to go into the technical discussion of verb movement here. Finite verbs in sub-clauses remain in final position; they don't move.

The *verb-second* parameter asks whether a language is a V2 language or not; there is a yes-no answer. (As in much of this discussion I am ignoring complexities; I hope that this is understood by linguists). Dutch is a verb-second language, as are many other languages around the world. English and French are *not* verb-second languages. Thus in order to answer whether children sometimes set parameters incorrectly (2a) we need to know whether, for example, they set the verb-second parameter correctly or not. The remarkable difference in grammatical structure that is related to the choice of finite or non-finite verbs has been a major tool in our ability to answer this and many other questions.

Optional Infinitives in Children

One of the major discoveries of the last decade in early linguistic development was the discovery of the *Optional Infinitive (OI) Stage* (Wexler 1990, 1992, 1994), which lasts in normal children from birth (so far as we can tell) to around 3;0.

(7) The properties of the OI stage are the following:
a. Root infinitives (non-finite verbs) are possible grammatical sentences for children in this stage
b. These infinitives co-exist with finite forms
c. The children nevertheless know the relevant grammatical Principles and have set their parameters correctly

(7a) tells us that young children often appear to leave tense out of their verbs which require it. For example, here are two examples from a young child (less than 3;0) speaking Dutch:

(8) pappa schoenen wassen
daddy shoes wash-INF
'Daddy wash (non-finite) shoes'

(9) ik pak ‘t op
I pick it up
'I pick (fin) it up'

The form of the verb in (8) (ending in *en*) indicates that it is a non-finite verb. Examples like (8) confirm (7a). But (8) is *finite*; it has a first person singular present tense, confirming (7b). Wexler (1990, 1992, 1994) and many other references analyzed individual subject data to show that at a particular age, the child produced *both* kinds of verbs, finite and non-finite[3].

But there is a crucial difference in the examples in (8) and (9). In (8) the verb (non-finite) appears in final position, where non-finite verbs go in Dutch. In (9), the verb (finite) appears in *second* position.[4] These examples are thus in accord

[3] Wexler (1994) suggested that the increasing proportions of finiteness with age made it natural to think that at extremely young ages children produce 100% non-finite forms. Wijnen (1998), DeJong (this volume) have produced evidence that this is so.
[4] (9b) is ambiguous between 2nd and final position, of course. I have included it to make the point that in counts of the finiteness/word order correlation, research on the OI stage has *not* counted the ambiguous forms like (9b) in deciding where the verb appeared. See Poeppel and Wexler (1993).

with (7c). The children are putting the finite verbs in second position, where they go, and they are putting the non-finite verbs in final position, where *they* go.

To show that these examples are not chosen arbitrarily it is necessary to count *all* the relevant verbs from children. Here is some data from a study of the development of 47 normally developing Dutch children (Wexler, Schaeffer and Bol in press).

(10)

Proportions of Optional Infinitives by age

age group	% OIs
1;07-2;00	83% (126/152)
2;01-2;06	64% (126/198)
2;07-3;00	23% (57/253)
3;01-3;07	7% (29/415)

For the youngest group, 83% of their main-clause verbs (i.e. 126 non-finite verbs out of a total of 152 verbs) are OI's, basically ungrammatical in the adult language. For the oldest group (3;1 to 3;7), the OI rate is only 7%. This is a well-documented trend in the study of OI's; the OI rate decreases over time. The same result holds in individual children; a child produces fewer and fewer OI's over time. Developing adult finiteness behavior (essentially 100% finite utterances) is thus not a question of learning at one time.

So it is quite clear that (7a) and (7b) hold of this population. To see that children produce both finite and non-finite utterances at a given age, individual children have to be studied. That in fact is the typical method of studying the OI stage, which for reasons of space I won't illustrate. I wanted to show what OI rates look like over a broad sample of children, so that the reader understands the great prevalence of OI's; to my knowledge there is no reason to think that any child in Dutch escapes the OI stage at the relevant age.

In order to test (7c), we have to see whether finite verbs appear in second position and non-finite verbs appear in final position. Wexler, Schaeffer and Bol did this calculation, following the usual procedure of only counting root (main) verbs, so as not to make the results look better by counting non-finite verbs that *should* be infinitival. The results are in (11), for the same set of 47 normal children, where only non-ambiguous order is counted.

(11)

Finiteness/position contingency normally developing children

all normal children	**V2**	**Vfinal**
Finite	1953 (99%)	11 (2%)
Non-finite	20 (1%)	606 (98%)

Almost 2,000 (99%)of the verbs in second position are finite; only 20 are non-finite. But more than 600 (98%) of the verbs in final position are non-finite; only 11 are finite. Finite and non-finite verbs could hardly be behaving more differently in terms of word order. Since most of the children are producing both finite and non-finite verbs, this lack of error also means that individual children are placing essentially all their finite verbs in second position and non-finite verbs in final position.

(7c) is supported in this data as strongly as anything in child development (or almost all of the cognitive sciences, in fact) ever is. Very little leeway has to be given to measurement error or noise, even at the youngest ages. This is the kind of data that psychologists studying cognitive development almost never see, close to categorical data. It seems quite reasonable to consider the small number of exceptions to the finiteness/word order correlation to be performance errors or some other kind of error of measurement.

Very-Early Parameter Setting, Learning and Imitation

From the earliest investigations of the OI stage (Wexler 1990ff.), data like this were taken to show that children set parameters correctly very early. In particular, this data shows that children set the V2 parameter correctly. From their earliest utterances, Dutch children place finite verbs in second position and they place non-finite verbs in final position. This is what would be expected if they knew that Dutch were a verb-second, verbal-final language and they produced OI's.

Children speaking V2 languages like Dutch and German not only place the finite verb in second position during the OI stage, but they place any major constituent in first position in a finite sentence, as is expected in a V2 language. Poeppel and Wexler (1993) showed data confirming this point in German, and there is a great deal of evidence that it is true.

On the other hand, children learning non-V2 languages, do not show the verb-second properties. They don't put finite verbs in second position (this can be seen, for example, in verb-final languages like Japanese and Korean), nor do they put any constituent into first position (for example, English-speaking children do not do this). In other words, while Dutch children show the behavior discussed above, children developing non-V2 languages do *not* show this behavior.

Wexler (1990ff.) argued on the basis of these kinds of phenomena that children set their verb-second parameter (*yes* or *no*, depending on the input language) correctly from the moment that the question could be asked, that is, from the moment that children entered the two-word stage, producing a verb and another constituent in the same utterance. (Before this stage, the question of correct parameter-setting can't be settled by production data because utterances of one word do not give word order information). He argued further that the same thing was true of all central parameters concerning clause structure and inflection -- see these papers plus Wexler (1998) for a discussion of several parameters. These parameters included the V2 parameter, the verb-to-tense (verb raising) parameter, word order parameters like VO or OV, and the null-subject parameter. So far as I know, there is no evidence that any parameter is mis-set by young children. This is the property I argued for in (12):

(12) *Very-Early Parameter-Setting (VEPS)*: From the earliest observable ages (around 18 months), children have set their parameters correctly.

It will take an advance in experimental techniques to determine whether VEPS is true at even younger ages. See Soderstrom, Wexler and Juczyk (2000) for some evidence in English that infant techniques might help us to settle that question of earlier ages.

One might question whether the strict correlation between word order and finiteness that we have shown actually does constitute evidence for correct parameter-setting. Perhaps children are only good imitators, perhaps they are a kind of imitating automaton that reproduces the input. Since children hear finite verbs in second position (in main clauses) and non-finite verbs in final position, perhaps they are simply repeating these verbs in the word order in which they hear them. Let's call this the automaton view (to distinguish it from a more sophisticated view of imitation which I will shortly discuss).There seems to be no way to maintain the automaton view, however.

First, the automaton view suggests that children don't actually understand sentences, they don't understand verbs and nouns and how to put these together, for example. This contradicts the experience of not only developmental psycholinguists, but also of parents. I can't possibly review the evidence here, but it is a distinctly surprising view.

Second, the automaton view doesn't explain why children always place finite verbs in second position, since they hear finite verbs at the end of clauses when these clauses are not main clauses. Somehow children would have to ignore these subordinate clauses. But how would an automaton that didn't analyze sentences know that a verb was part of a subordinate clause?

Third, the automaton view doesn't have a learning theory, so far as I know. Note that it's not enough for a child to be able to learn that one form follows another form. To capture even simple V2 facts, the child will have to associate the finite verb (presumably a verb with a certain phonology for this view) with "second" position, and the non-finite verb with "final" position. In addition the child will have to know what counts as a constituent, and that any constituent (including adverbs) can appear in "first" position. That is "second position" is *not* defined as "second word in a sentence." (We know that the child produces utterances in line with this knowledge).

Fourth, the automaton view doesn't explain why children produce OI's at all. Since basically all simple sentences in the input are tensed, why does an imitating child go out of her way to produce untensed sentences, sentences quite at odds with what has been produced?

Fifth, why does the child in many languages produce such a large percentage of OI's at an early age? Notice that 83% OI rate for the 1;7 to 2;0 children in (10). Wijnen (1999) has argued that at the very earliest ages in Dutch, there are actually 100% OI's. If the child is imitating the input, even if some kind of stray input or mis-analysis led to the occasional utterance of an OI, why should almost all the child's early utterance be OI's, which are not attested in the input.(See for example Poeppel and Wexler (1993), who found no input OI's). Even if parents actually use a few OI's, for whatever reason, why should a young child's productions be overwhelming OI's? This behavior is quite the opposite of imitative behavior.

Sixth, and quite strikingly, there are a number of systematic errors that children make that have no basis at all in the input, but which relate to their understanding of OI's as non-finite. A major example are the errors on subject case in English that do not show up in Dutch or German. We will return to a discussion of these errors and why they are so difficult for an automaton model to handle in a succeeding section.

This is just a beginning of a set of questions that the automaton view seems quite incapable of dealing with. Recently there has been an attempt to make the Imitating view more sophisticated, to continue to think of the child as having no linguistic knowledge, but of having a richer set of learning mechanisms than the simplest behaviorist views would have allowed. A prime example of such a theory is that of Tomasello (2000).

Tomasello argues that young children have essentially no knowledge of linguistic categories, principles or processes. He writes (p. 241) that he posits "...that in the beginning children make virtually no linguistic abstractions at all (beyond something like 'concrete nominal')" and (p. 247) "...that at younger ages children simply do not possess the abstract syntactic competence characteristic of older children and adults." Although he may not specify ages exactly clearly, the surrounding discussion suggests that he is claiming that children until about age 3;0 don't have linguistic categories. The category that Tomasello concentrates mostly on is the category of "transitive verb." He is claiming that children until about 3 don't even have the category of transitive verb[5]. Although he doesn't discuss these processes, his theory would assert that young children don't have

[5]Critiquing Tomasello's' supposed evidence for his view that two-year-olds don't even know the category of transitive verb is beyond the purpose of this paper. Let me just point out that Tomasello's experiments which show that children don't much generalize from inchoative novel verbs (*the ball is meeking*) to transitive novel verbs (*the boy is meeking the ball*) in no way establishes that children don't have the concept of transitive verb, contrary to Tomasello's claims. For this is not a systematic syntactic pattern in English (*the book fell/*the boy fell the book*), and it wouldn't be a good generalization for the child to draw; it's not a generalization supported by universal grammar at all. As for the experiments that show that young children who are taught novel verbs in passive form don't reproduce them in active form, the result is no surprise to those generative accounts (Borer and Wexler 1987, 1992, Babyonyshev et al 2001) which say that children at this young age don't have the syntactic basis for verbal passives; the linguistic system hasn't sufficiently matured, that is, the A-chain deficit theory. A more telling experiment (on the assumption that the novel verb technique is tapping children's linguistic abilities at all) would be to teach the children the novel verb in passive form, (*the dog was meeked by the cat.)*, then ask a passive-inducing question (*what is happening to the dog)?* . Especially if the verb were *non-actional* (Maratsos et al 1983) (and thus had no homophonous adjectival passive, the A-chain deficit theory would predict that the child couldn't answer with a passive form despite the passive introduction of the verb (see Borer and Wexler 1987, Fox and Grodzinsky 1998, Babyonyshev et al 2001). But this experimental type wasn't done by Tomasello.

such processes as verb movement or noun phrase movement until (if ever[6]) a much older age.

Tomasello claims that the classic learnability arguments have been made against learning theories that are only "straw men" --"simple association and binding induction" (p. 247)[7]. He claims that these arguments don't hold if these straw men are replaced by "...the more cognitively sophisticated learning and abstraction processes involved in intention reading, cultural learning, analogy making, and structure combining."

Tomasello's description of these "more sophisticated" learning processes is not clear enough to see how they would actually work; there is no attempt at formalization, and nothing in the way he describes them makes them look new or sophisticated in any particular way.

But it is worth considering the most important process that Tomasello mentions and that he mostly discusses, namely "intention reading." Tomasello agrees with generative-based critiques that classical imitation "very likely plays only a minor role in language acquisition." Tomasello saves the imitation theory by renaming imitation; he calls it "mimicking" (p. 218). Then Tomasello uses the name "imitation" for a completely different process, one in which the learner understands the intention of an actor and tries to reproduce the intention. "In cultural (imitative) learning, as opposed to simple mimicking, the learner understands the purpose or function of the behavior she is reproducing." (p. 238). "Thus, a child might hear her father exclaim, "Look! A clown!" To fully understand his linguistic behavior (with an eye toward reproducing it) she must understand that her father intends that she share attention with him to a particular

---

[6]I write "if ever" because Tomasello would really like to argue that even adults don't have such processes; the sub-text (often explicit) of his paper is that linguistic theory is not describing psychologically true phenomena. There is no space here to illuminate Tomasello's misunderstandings. He somehow thinks that linguistic theory is concerned with mathematical rather than psychological properties; he seems to not understand that mathematics is a tool used scientifically to describe scientific theories, as in ,for example, physics. On Tomasello's reasoning, the theory of physics would not be "physical", since it uses mathematics. I urge the reader to read Tomasello's paper, to see that this seems to be his reasoning. There seems to be a kind of tradition in parts of psychology that says to attempt to understand language and its development in precise, scientific terms is somehow wrong, that language can't be studied like other fields of science, it just isn't precise enough. Tomasello's attempts seem to fall into this category. It is difficult to square the incredible regularity and interaction of phenomena that I have reported in the text with the anti-precise notions of language and its development that Tomasello seems to be urging.

[7]This claim is false. The classic learnability arguments were made assuming *any* mechanically specifiable (i.e. actually computable in a well-grounded, accepted sense in the cognitive sciences) learning theory. See for example, Wexler and Hamburger (1973), Wexler and Culicover (1980). It was shown that for certain classes of linguistically-motivated processes, no specifiable learning theory (that didn't assume specific linguistic knowledge) could learn all the possibilities. Nothing special had to be assumed about "simple association." As for the denial of "blind induction", this is exactly what the theories that he is attacking do.

object; that is to say, understanding a communicative intention means understanding precisely how another person intends to manipulate your attention" (p. 238). "To comprehend the totality of an adult's utterance, the child must understand his overall communicative intention and also the communicative role being played by the various constituents of the utterance." (p. 239).

The theory seems to be this: the child can somehow (unspecified) figure out what the adult intends to say, and the child then maps the string of sounds to this intention; furthermore in some even more complicated (unspecified) way, the child can figure out what the *constituents* of the intention are and how these are mapped to the *constituents* of the sentence. (see p. 239).

This suggestion appears to be the major theoretical proposal that Tomasello makes concerning how language learning takes place. Interestingly, the basic assumption -- that children need to be able to figure out something about the intended meaning of an utterance in order for language learning to proceed -- is a staple of generative acquisition theories, at least since Wexler and Hamburger (1973), Hamburger and Wexler (1973) and Wexler and Culicover (1980). Those authors argued that "semantic information" had to be available to the language learner, and they gave an explicit discussion of this assumption, and what had to be assumed to make it work. (See Wexler (1978) for particular attention to this point). Essentially the semantic information helped the learner to construct the "deep structure" of the sentence.

The argument was mathematical and empirical, in the tradition of scientific reasoning. Namely, certain linguistic variation possibilities couldn't be learned if only "surface information" (Wexler and Hamburger's term) were available to the learner; the learner had to supplement this with information concerning the intended meaning.

In more modern learning theories, changed as the result of more recent discoveries concerning the form of syntactic parameters, the same basic assumption about the necessity for semantic information helping the learner is made. Thus Gibson and Wexler (1993), for example, assume that semantic information helps the learner to figure out which NP is the subject of the sentence and which is the object.

The essential point is that Wexler and Hamburger and Wexler and Culicover showed that even *with* the semantic information, there were unlearnable linguistic processes unless it were assumed that the learner had access to grammatical universals. In the case of Wexler and Culicover, they were very concerned with showing that learners could learn their language from fairly simple sentences, since very young children can't handle very long sentences. They demonstrated that transformational grammar (in a specified sense) could be learned from sentences with no more than two degrees of embedding, so long as children had access to semantic information *and* universals of grammar. *Both* were necessary. And this was formally, mathematically demonstrated.

Although particular theories have changed, this is the essence of theories of language learning in the generative tradition -- both semantic information and grammatical universals are necessary for language learning. Tomasello posits that semantic information (what he has relabelled "intention reading") is necessary and helpful for language learning. I agree completely and am glad that he has accepted these arguments from generative learning theory. But without making any arguments, Tomasello also says that semantic information (intention reading) is *sufficient* for language learning. There are good arguments otherwise, and Tomasello makes no counter-arguments.

Let us call Tomasello's theory the "intention learning" theory (with its most singular characteristic being that the child has no grammatical categories). How does the intention learning theory do on the six problems that I mentioned above for the automaton view? It seems to overcome the first problem, since, unlike the automaton view, the intention learning theory does assume that children attempt to understand sentences. But on problems 2 through 6, I see nothing in the intention learning theory that can solve these problems. If the child doesn't have access to grammatical categories or to the setting of parameters, there is no way to explain the patterns of OI behavior.

In fact, Tomasello *does* explicitly try to explain OI behavior via the intention learning theory. He writes (p. 240), that "...a major part of the explanation is very likely the large number of non-finite verbs that children hear in various constructions in the language addressed to them, especially in questions such as *Should he open it?* and *Does she eat grapes?* The child might then later say, in partially imitative fashion: *He open it* and *She eat grapes*."

This seems to be an attempt to deal with one of the problems I raised, namely problem 4: why does the child use OI's at all? Tomasello here retreats from the "intention learning" theory and move back to the "mimicking" theory, which he earlier rejected. For the input was in the form of a question, on his account, and yet the child uses the forms in a statement. Surely the difference between question and statement is one of the most simple and basic aspects of communicative "intention." Any intention-reading learner would and should pay attention to the major difference between the intentions of questions and statements, and wouldn't associate a form that goes with one intention (the question) with another intention (the statement). And there is good evidence that the child *does* pay attention to this difference; young children do not use auxiliary-first (inverted) order to make a statement; they *would* do that if they imitated question word order when they were making a statement. What does intention-reading theory have to say about why word order isn't mis-learned whereas verb form *is* mis-learned?

So Tomasello seems to reject intention reading here, and goes back to mimicking. But then we have all the problems associated with mimicking, that Tomasello explicitly acknowledges, as he rejected mimicking.

But putting aside the fact that Tomasello's suggestion about why OI's exist contradicts his theory (being more allied with the automaton view), we still have many problems. In most of the other OI languages discussed in the literature (e.g. all the Germanic languages except English), questions are not typically asked by using a finite inverted auxiliary, plus a non-finite verb; they only are when there is a modal or other type of auxiliary in the meaning of the sentence. But for a main verb, the verb itself is used: *isst sie Ei/eats she eggs/'does she eat eggs'?* (German). Children tend to use 100% (or almost 100%) OI's in their youngest ages. On Tomasello's proposal, this means that children use the question model for the form of the verb almost 100% of the time, for some unspecified reason ignoring the declarative input. But then children should produce the finite verb before the subject almost 100% of the time, when they're using only a finite verb, since in questions the main finite verb always precedes the question. But they very rarely do this, certainly nothing like 100% of the time.

None of questions 2 through 6 are answered by Tomasello's suggestion. For example, consider question 2; why don't children use the input that has finite verbs at the end of the clause (any sentence with a finite verb in an embedded clause) to imitate and thus put finite verbs in sentence final position instead of verb second position. We have already seen that they essentially never do this. So children use some misleading input to lead to an almost 100% error rate but don't use other misleading input at all, they get a 0% error rate. The theoretical and technical tools don't seem to be available in intention reading learning theory (or in the mimicking theory to which Tomasello retreats) to explain substantial empirical properties of development.

It is very important to reiterate that I am not criticizing the notion of intention reading/leaning; although a very difficult concept to work out explicitly I have long argued that it is part of language learning and the generative field mostly accepts this proposal. So it is useful for Tomasello to update the multi-purpose learning school of language learning so as to help it to become more cognitive, at least recognizing the need for the child to attempt to understand what is being said. This is an advance beyond automaton theories. What I *do* find wrong in Tomasello's theory, and other anti-generative, anti-nativist theories of its type, is the claim that genetically-guided knowledge is *not* part of the child's endowment. I welcome Tomasello's recognition that traditional learning theory approaches were too limited, and hope that as he attempts to add concepts to the theory he will keep adding the ones that have been proposed in generative-based theories. If he and others attempt to actually work out process models of learning, as has been carried out in detail in generative learning theories, he might discover that intention learning is not sufficient.

I will return to intention learning when we discuss case errors (problem 6). Meanwhile we can conclude that the strict patterns of morphology and word-order correlations that children produce in the OI stage is good evidence for their having set parameters correctly.

OI's in English

One of the reasons that it took so long to discover the OI stage is that so much modern work (since roughly Brown (1973) on language acquisition has been carried out in English or influenced by research on English. Unlike other Germanic languages, Romance languages and many/most other languages that have non-finite forms, the English infinitival verb has no audible inflection. The infinitival form of the verb sounds just like the stem of the verb, *to go*, *to walk*, *to eat*. Compare the infinitival form of *speak*:

(13)	French:	parl	+	er
	German:	sprech	+	en
	English:	speak	+	Ø

The English infinitival suffix is phonetically zero, it's unpronounced. Therefore OI's weren't discovered because there was no obvious "extra" morpheme that had been added to the verb while the tensed/agreement morpheme was omitted. In Dutch we saw that *en* is added to the stem when children produce an OI, and this form is extremely noticeable, it's clearly not a stem, it doesn't belong there. So it's just an accident of English that OI's are a less obvious phenomenon.

But OI's clearly exist in English, as Wexler (1990ff.) showed. As was well-known since Brown (1973), children often produce what sounds like the stem form instead of the third person singular form, for example *push* instead of *pushes*. Similarly they produce what sounds like the stem form instead of the past tense form, for example *push* instead of *pushed*. Wexler argued that these forms were expected if Tense were omitted from the structure and a non-finite form was therefore the appropriate form. The analysis of the form would be as in (14), where the phonetically empty morpheme Ø is the spell-out of the non-finite/infinitival morpheme in English:

(14)	a.	pushes	--->	push	+	Ø
	b.	pushed	--->	push	+	Ø

These non-finite forms in English showed all the properties of OI's. For example, the proportion of non-finite English forms decreased in a child as the child aged, just as OI's do. Wexler also showed that children understood the grammatical properties of the tense morphemes in English, just as they understand the properties of the finiteness morphemes in other languages. For example, they appear only in the correct positions, children for example don't say *Mary not pushes the chair*. If they omit the auxiliary they say *Mary not push the chair*. Furthermore, children understand the semantic properties of the tense morphemes; they don't use the present tense morpheme when past is appropriate (**pushed* ---> *pushes)* or the past tense morpheme when present is appropriate[8] (*pushes* --->*pushed*) (Rice, Wexler and Cleave 1995, Rice and Wexler 1996
for children 3 and above, Schütze and Wexler 2000 for 2;6 and older). In either case, children in the OI stage might use the "stem" (non-finite) form instead of the correct tense form, but they won't substitute the wrong tense form. All these predictions and many others follow from the assumption that young English-speaking children are in the OI stage even though there is no obvious "infinitival" morpheme.

[8]These experiments were done in ordinary discourse contexts, and the prediction of course is just for these. Children might have somewhat different properties of tensing in special contexts, for example in narratives children even at older ages sometimes use present tense more than is used by an adult. But these differences involve conventions of discourse; children clearly aren't making mistakes on whether to use a particular morpheme as a present tense or past tense morpheme.

The OI stage is called that, the Optional stage, because of its most prominent characteristic in the original languages in which it was discovered. But as Wexler (1990ff.) argued, the stage is in no way limited to what are traditionally called "infinitival" forms. Rather, the prediction is that *non-finite* forms occur; often these do not take the form of infinitives. For example, in many sentences in English, finiteness is marked only by an auxiliary. These auxiliaries have no semantic function other than to mark the inflectional properties of finiteness (tense and agreement). The OI stage predicts that these morphemes are omitted for the same reasons (which we have not discussed yet) as are the inflectional finiteness morphemes on the main verb. This prediction is strongly confirmed, in English and in many other languages. An English example is that auxiliary *be* is quite often omitted by children: *Mary going*.

The prediction is quite strong. Namely, when measured by rate of use in obligatory context, the finiteness morphemes in English should pattern together in development, taking a very similar course, showing relatively minor fluctuations from each other. Exactly this was shown by a detailed analysis of longitudinal data from a large group of children using structural equation modeling in Rice, Wexler and Hershberger (1998).

At the same time, other morphemes that share identical surface (phonological) patterns don't behave similarly to the finiteness morphemes. As Rice and Wexler (1996) show, plural *s* in no way patterns like third person singular *s*. The latter is a finiteness (tense) morpheme, and thus part of the OI stage predictions, the former isn't. Plural *s* develops much faster than third person singular *s*; there is hardly any delay. It does not pattern w/ the finiteness morphemes, as expected. For this any many other reasons, we know that the use of morphemes that we are discussing is not delayed because of their particular surface or phonological properties; rather there is a deeper grammatical factor that underlies the OI stage.

Subject Case

One of the most important features of the OI stage analysis is that it allows us to bring together in the same system a myriad of phenomena which have been known to some degree but which have previously had to be understood in completely different terms. It has been known for a very long time that children in English often substitute ACCusative case pronouns for NOMinative case pronouns. They often produce forms like (15):

(15) a.her going

b. me here
c. him like candy

In English, subjects of root clauses are NOM, *he*, *I*, *he* instead of the ACC forms used in (15). Schütze and Wexler (1996) showed that the case errors that children made were always substitutions of the ACC form for the NOM form; they essentially never substituted NOM for ACC. That is, in object positions, for example, children always used ACC forms; we don't hear children produce utterances like *Mary likes he*.

So one of the major facts that has to be explained is this asymmetry. It can't follow from any kind of standard "frequency" argument. As Schütze and Wexler point out, Colin Philips has shown that in the input NOM forms are much more likely to appear than are ACC forms. So the children are going out of their way to substitute the form that is far *less* frequent in the input.

Another major fact is that the incorrect ACC subject forms like (15) essentially never appear when the verb is finite. They only appear when the subject is an OI. Here, for example, is the table from Schütze and Wexler (1996) analyzing the CHILDES data of Nina (McWhinney and Snow 1985). (See also Loeb and Leonard 1991)

(16)

Nina's 3rd person singular subject pronouns: finiteness vs. case

Subject	Finite Verb	Nonfinite Verb
he+she	255	139
him+her	14	120
Percent non-NOM	5%	46%

There is an extremely small possibility of using ACC subject pronouns with finite verbs.[9] Schütze and Wexler provide statistical arguments that this effect is not one simply of correct case and correct finiteness developing simultaneously; rather a

[9] The counter-examples are actually smaller than 5% because for independent reasons Schutze and Wexler develop a more complicated model, the AGR/TNS Omission Model, or ATOM, which is a better description of the OI stage than the Tense Omission model that I am essentially discussing here. Under ATOM, some of the 5% counterexamples are not counterexamples. I briefly describe ATOM later in the text, but I won't take the space here to explain the particular expectations about subject case.

child at a given age shows a strong correlation between finiteness and case-marking such that the child will alternate finite and non-finite verbs, and NOM and ACC subject case for pronouns, but will never use ACC case with a finite verb: **her is going now*.

The empirical linkage of ACC subjects to OI's suggests that in fact ACC subjects are only possible *because* OI's exist. In my earlier work (e.g. Bromberg and Wexler 1995) I suggested that ACC pronouns were default pronouns, used because TENSE was missing in OI's, on the assumption that NOM was only possible with tensed subjects. Schütze and Wexler provided a more detailed model, arguing that it was the AGREEMENT part of the finite verb that licensed NOM case. Since OI's lacked agreement, they couldn't license NOM case on the subject, and the default pronoun was used. There is good reason to believe that in fact, agreement is responsible for NOM, but I won't take the space here to go through the arguments. See Schütze and Wexler's paper.

Schütze and Wexler proposed the AGR/TNS Omission Model (ATOM), which says that in the OI stage, either AGR or TNS is optionally omitted by the child. The non-finite form of the verb is used whenever either AGR or TNS is missing. When AGR us present, whether or not TNS is missing, the NOM subject pronoun is selected. When AGR is missing, even though TNS is present, the default case form of the pronoun is selected.

What is a default form of case? It is the case form that is used when there is no structural case position. For example, in English, we say *it's him*, not **it's he*. Or we answer the question *who wants candy?* with *me*, but not with **I*. In these positions of the pronoun, nothing in the structure of the sentence dictates whether the case should be NOM or ACC, so the default form takes over, and this is ACC in English. Schütze and Wexler in fact showed that English-speaking children in the OI stage always correctly used the ACC form of the pronoun in true (adult) default positions.

As Schütze and Wexler discuss, in German and Dutch the default case of noun phrases is NOM, not ACC. And they discuss the literature which shows that the English subject case error is *not* replicated in German or Dutch. This is exactly as predicted. When the verb is an OI in German or Dutch, the child will use the default form, just as in English. But the German/Dutch default form is NOM, so the child will use the NOM form. And this is exactly what happens. In contrast to the 46% ACC rate for Nina above, German or Dutch children in the OI stage use essentially no ACC subjects, even of OI's. The rate is almost 0%.

So now we have another fact that children in the OI age-range have learned. They know what the default form of case is in their language. The default form varies from language to language: ACC in English, NOM in German/Dutch. This means that it must be *learned* from experience. Given the results above, we know that children in the OI stage have correctly learned the default form in their language, even when these have opposite values, e.g. ACC in English, NOM in German/Dutch. So just as in the case of parameters (and default case could be looked at as a kind of parameter, though it doesn't have to be) children learn the language-specific aspects of simple clause and inflectional structure very early and very well.

It is no mystery how children learn the default forms. Although we don't know the answer, because nothing is directly known about learning in language (because of the difficulty in observing an act of learning), it seems pretty reasonable to infer that children choose as the default form just that form that appears in "default" contexts, that is contexts where there is no structural case position. Given their knowledge of the Principles of UG, children can calculate which contexts these are, and it remains only to learn which form appears in these contexts. This learning is done by simple observation, given the calculations that children perform.

So the Principles and Parameters framework, together with the theory of Optional Infinitives understands why children behave as they do, why they give these complicated and specific interactions between tense and case, for example. Furthermore, this theory understands how children could easily learn default case.

How would an Imitating/Automaton model attempt to deal with these facts? Since ACC forms don't occur in subject positions for the most part, why does the child produce such? For the Intention Learning version of an imitation model, Tomasello (2000, p. 240) suggests that children in English imitate the kind of pronoun they hear in constructions like *let her open it*; "they may just imitatively learn the end part of the sentence..." These are small clause constructions, which take ACC subjects. This means that children have to ignore the fact that these forms never occur as the first word (subject) of the main sentence. One wonders how they could ignore this fact and at the same time learn, say, the verb-second property of German or Dutch, which they know so exquisitely? As a learning theorist, I would be delighted to see a learning mechanism that could have both those properties.

Tomasello says that children "basically **never**" use NOM pronouns for ACC pronouns[10] (**Mary hit I*) and that the reason for this is that "...they never hear adults say anything like this in any linguistic construction." What he must mean is that NOM pronouns never follow verbs, that is, he is assuming that sequences of words occurring next to each other are crucial for imitation, although he doesn't state his assumptions explicitly. At any rate, his claim is false: consider sentences like *Mary knows I like* candy, or *who did Mary tell I like candy?* In the first sentence the NOM pronoun *I* follows the verbs *knows*; of course, *I* is not the object of *knows*. In the second sentence, the NOM pronoun *I* follows the verb *tell*. Of course, the pronoun is not the *object* of *tell*; in syntactic terms there is a trace of the object between *tell* and *I*. But Tomasello is assuming that children have no knowledge of such syntactic categories, of relations like subject and object, of traces; presumably they are only paying attention to sequences of words. So there is evidence in the input for NOM pronouns following verbs, in the sense of input revidence relevant to Tomasello's model.

But the situation is far worse. For we know that in German or Dutch children do not use ACC subject pronouns Yet the German or Dutch equivalents of *Mary saw him go* exist, with ACC NP's as the subjects of the small clauses. So the input situation in German or Dutch is similar to the input situation in English, with respect to the juxtaposition of ACC case and non-finite verbs (*him go)*. On the imitation learning model, Dutch and German children should produce as many ACC pronoun subjects as English-speaking children. But they don't produce any.

The methodological problem with the imitation learning view is that its mechanisms are not well-specified; each time a phenomenon in children is discovered the model can make up a reason why there is evidence in the input for it. This is what Tomasello (p.232) called a "fudge factor" when he discussed maturation. But maturational theories (Borer and Wexler 1987, 1992, Babyonyshev et al 2001 among others) make cross-linguistic predictions about differences in development, which could easily invalidate the model. Tomasello hasn't tested his ideas against cross-linguistically different predictions -- so far as I know the observations that I have just made are the first such tests of the imitation learning ideas, tests which have a negative outcome. In contrast to formal maturational ideas that have been proposed, it is harder to make such predictions for the imitation learning view, because what counts as an adult model and what

[10]Tomasello doesn't reference any papers on the pronoun facts in children, but the patterns he is assuming are some of those argued for in Schutze and Wexler (1996). He treats the pronoun facts in the same paragraph as the use of OI's, suggesting that he implicitly recognizes that it has been argued that the phenomena go together in the OI stage.

counts as imitation and what counts as an intention have not been specified sufficiently. Nevertheless, I believe that the arguments that I have just made pretty much show that the imitation learning view is wrong for the cases that I have discussed.

The theory that I have proposed, linking Principles, Parameters, and OI's explains this result; it predicts it clearly. Children *learn* the default case form of their language. They can't learn this from the subjects of OI's, since these sentences don't exist in the adult input. But they learn the default form, as I suggested above, from sentences with NP's that are not in a structural case position. Once they learn the default case form, they use it for the subject of OI's.

But the imitation learning view has no recourse to a notion like "default" case. For such a notion presupposes a notion of "structural" case. The default case is just the case that is used when the NP is not in a structural case position. The imitation learning view by definition asserts that the child has no implicit notion of structural case. Thus it can have no implicit notion of default case, in the relevant sense[11].

Note that what an imitation learning view --like all such views which deny that young children have any kind of computational linguistic system -- would like to assert is that the notion of *default* that I have defined can be replaced with a notion based on frequency in the input. That is, it would like the default form to be the most frequent form in the input. But this is just false; we have already seen that the NOM form in English, which is *not* the default form, is by far the most frequent in the input. In Dutch and German, the default form *is* NOM; although I haven't seen any data, presumably NOM is the most frequent form in the input in these languages also. So the sense of default that is needed is orthogonal to frequency in the input. What is needed is a computational notion of default, part of

[11]Tomasello's paper appears to not understand the relationship between finiteness and subject case, a classic fact about languages that any theory would have to take account of. He makes a point of discussing what he calls the "incredulity construction" (p 236), examples like *My mother ride a motorcycle!* He writes that this construction "is very odd from the point of view of the majority of English sentence-level constructions because the subject is in the accusative case ...and the verb is non-finite." He somehow wants to remove this construction from the "core" of the language. He seems unaware of the fact that the construction has been discussed (sometimes at length) in many OI papers. And unaware that the ACC case *follows* from the non-finiteness of the main verb. For example, in languages with NOM default case, the subjects of this construction are NOM, despite the non-finite verb. For Tomasello, the construction is just some strange thing that doesn't obey grammatical rules-- he thinks it's special to English. But its ubiquity and law-like behavior make it understandable within UG analyses.

the child's system of language. Children, it turns out, and not surprisingly, *have* a computational system of language.[12].

Variation across Languages in the OI-Stage: the NS/OI Correlation

One of the most interesting facts in the study of OI's is that, although many languages go through such a stage, many do not. For example, Italian, Spanish and Catalan do not go through the OI stage. The percentage of OI's in these languages, even at very early ages, is extremely small. More than 20 languages have been studied at this point, and there is a generalization that fits the data perfectly so far, as in (17) (Wexler 1998, also see Sano and Hyams 1994):

(17) The Null-Subject/Optional Infinitive Generalization (NS/OI): A child learning a language goes through the OI stage only if the language is *not* an INFL-licensed null-subject language.

NS/OI says that Italian, Spanish and Catalan won't go through the OI stage because they are null-subject languages. German, Dutch, English and French, on the other hand, are not null-subject languages, and they do go through the OI stage. See Wexler (1998) for a discussion of more languages and more data.

The question is, why should NS/OI hold? Wexler (1998) derives the existence of the OI stage as well as NS/OI from the assumption that what characterizes young children is a particular limitation on their computational systems, the *Unique Checking Constraint*.

(18) *Unique Checking Constraint (UCC)*: Children can only check once against the D-feature (the Determiner feature, that is, the feature that characterizes noun phrases= NP’s) of their subjects, whereas adults can do this more than once.

UCC is a developmental constraint on the computational system of language; it holds of young children and fades out over time -- it is *not* a constraint on the adult grammar (UG). Moreover, UCC is not subject to parametric variation, it isn't that some adult languages have UCC and others don't and that the child has to *learn* whether UCC holds. UCC is simply a constraint on children at a particular

[12]In a way, the intention learning model's lack of specification of a learning theory is in line with the historical foundations of such theories in Psychology. The most famous (radical) behaviorist of all -- B.F. Skinner -- wrote a famous article in which he argued that theories of learning are not necessary. This scientific primitivism is one of the reasons that Skinner's research, despite its many interesting experimental findings on schedules of reinforcement in rats, has had little impact in modern cognitive science.

immature time. One might think of it as parallel to constraints which don't allow children to walk at a particular time.

I just wanted to briefly discuss the UCC here, to give the character of the explanation; see Wexler (1998) for a full discussion of UCC. Let me try to explain here in an intuitive manner how UCC works to predict NS/OI. First, as I mentioned earlier, Schütze and Wexler (1996) argued that the OI stage is best described by the AGR/TNS Omission Model (ATOM). There are two inflectional functional categories, AGR (Agreement) and TNS (Tense). ATOM says that in the OI stage, either AGR or TNS is omitted by the child. This yields OI's, since many inflectional morphemes on verbs can't be inserted without both AGR and TNS being present, the result being the infinitival morpheme, thus the OI. For example, *s* in English specifies both agreement (third person singular) and tense (present). If AGR or TNS is omitted, *s* can't be inserted, and the non-finite morpheme (the phonetically empty morpheme in the case of English), *en* in Dutch, etc.) is inserted on the verb. Schütze and Wexler argue for ATOM on the basis of the particular constellation of effects of subject case errors.

Why does ATOM hold? Wexler (1998) argues that ATOM follows from UCC. In current syntactic theory (Chomsky 1995), it is argued that functional categories like AGR and TNS have D-features, and that these D-features (unlike the D-features of NP's) are uninterpretable. Therefore, to obtain a coherent meaning, the uninterpretable features must be eliminated. This is done by checking the uninterpretable D-features of AGR and TNS against the D-feature of a NP (the subject NP). The idea is that a subject NP has to check the D-features of both AGR and TNS. For the child, UCC prevents this from happening. Therefore children omit AGR or TNS. Thus the UCC implies that the OI stage exists and is described by ATOM. In other words, the OI stage results from the difficulty in the child's computational system of checking some syntactic features.

Informally, one might think that UCC prevents a subject from moving to both AGR and TNS, and that if a subject hasn't moved to these functional categories, they are ill-formed: all verbal functional categories demand to see a near-by subject. Thus the child has to eliminate either AGR or TNS, so as to make the sentence well-formed. The child's grammar, like adult grammars, won't tolerate AGR or TNS without a local subject.

An Italian-speaking child should have the same difficulties due to UCC, and in general she does. However, in Italian, AGR doesn't have to be checked against, because it itself is interpretable, as the subject of the sentence (the traditional idea about null-subject languages). Thus the subject NP in Italian only has to check against TNS, not AGR, and this amount of checking is not too much, it doesn't violate UCC. So the Italian-speaking child (or a child learning any null-subject language) does not have to omit AGR or TNS in order to satisfy the UCC. Thus all features are specified in productions of the Italian-speaking child; there is no OI stage. This is why NS/OI holds.

Using the same informal analogy, we can say that the subject only has to move to TNS in Italian, not to AGR at all (there is grammatical evidence that this is correct). This is because AGR itself operates like a subject in null-subject languages. Since only one movement is necessary, UCC isn't violated; the child has no reason to omit either AGR or TNS, and keeps both. Thus the finite morpheme, which depends on both AGR and TNS, can be inserted, and the child doesn't produce an OI.

UCC will still have an effect on children speaking languages like Italian, but they won't produce main clause infinitives, for reasons that I've just given. However, they are predicted to omit auxiliaries (for reasons given in Wexler 1998), and they do, during the OI age-range.

Cross-Linguistic Variation in Development

Although the description in the last section of the underlying theory of the OI stage (the UCC) was very brief, I wanted to introduce it just so that I could discuss the character of the explanation. We still have to ask why UCC holds and how it goes away as children age. But given UCC as a constraint on young children, we not only see why the OI stage exists, but we also see why many languages do not go through the OI stage.

In other words, there is an interaction between developing principles of the computational system of language (e.g. the UCC) and the actual language that the child is learning. Since the child is such an excellent learner of parameters (VEPS; also see Wexler 1998 for arguments that the child learns correctly from very early whether or not her language is a null-subject language), she knows whether or not her language is a null-subject language, that is, whether or not AGR has to be checked, and therefore UCC doesn't come into play in a language where the child isn't checking AGR. It is the *interaction* of universal developing principles and what is learned in a particular language that determines the linguistic behavior that we observe.

Note that the model does *not* say that children learn the parameters of Italian better than they learn the parameters of English or Dutch. The relevant parameters are all learned quickly and well (VEPS). It just turns out that once the child has learned the parameter values in different languages, these values interact differently with the universal developing principles, that both English and Dutch-speaking children are subject to (e.g. UCC).

This is quite a different picture than the traditional one in generative grammar-oriented studies of linguistic development, and also of more traditional studies. It is a picture which assumes (and shows) that children are excellent learners of language-particular properties of language. But there are some universal constraints on the developing child that might not exist on the adult, and these interact with the principles to form what looks like very different behavior. But in now way or it is *learning* deficit. After all, the detailed learning (of parameter values, of default case, of agreement and other inflectional forms) is exquisitely precise. Children are excellent learners of language-particular facts and they know universal grammar principles. However, they have some particular computational limitations as a result of their immaturity.

In fact, there are a wide variety of developmental differences across languages that the OI model that I have outlined explains. Some of these effects are quite interestingly subtle; for example, effects on rates of OI's, as opposed to the presence or absence of OI's. In particular, there are large differences in OI rates across languages that *do* have OI's in the appropriate age-range. These effects are understood by the interaction of the particular morphology of the language with the ATOM, which describes the OI stage. For example, English children in a particular age-range show a larger rate of OI's than do Dutch children at the same age. These results are understood by an analysis of the verbal morphology of the 2 languages, and the application of ATOM to this morphology. The differential rate is predicted. See Wexler, Schaeffer and Bol (in press) for the analysis and data. There are quite a few other cases like this, which we have no space to discuss.

In recent research the UCC has been applied to explain an even more diverse range of phenomena in the OI age-range. Thus Hagstrom (2000) and Baek and Wexler (2000) have explained a particular well-known word order error in the development of Korean using the UCC. Namely, in the so-called "short form negation" structures, *an/not* normally appears after the object, yielding the word order in (19a) in this SOV language:

(19) a. Subject Object an V

b. *Subject an Object V (child form)

Children, however, often produce the form in (19b). Although well-documented as an existing error, there has been no satisfactory explanation of why children go strongly against the input and create the wrong form. Hagstrom and Baek and Wexler propose that in adult Korean, the object raises, checking twice. Thus the UCC prevents this second checking, forcing the object to remain in a lower position, and thereby creating the word order error in (19b).

Baek and Wexler show that a predicted correlation holds, namely that when the child fails to raise the object (19b), he never inserts ACC case, although he often inserts ACC case when he does raise the object (19a). There are a number of other phenomena that are predicted and tested, and a detailed syntactic theory is given. The point is that a constraint on child grammar that explains the OI stage (and the failure of the OI stage to hold in some languages) also explains a completely different type of error in an unrelated language. Developing constraints have effects throughout the grammar. What looks to be unrelated phenomenologically is in fact the result of the same cause. We don't see a (phenomenologically) comparable kind of error in English because English doesn't have the same double-checking process of object raising.

The implications of the method and results are striking. It is a truism of research in developmental psycholinguistics that children's behavior looks quite different in different languages. Of course, we expect that different developing languages will exhibit properties that are different simply because the languages themselves differ. But the errors look different too. The general problem in the field is very old and it had been hard to figure out its solution. Why should children subject to universal principles make a different kind of error, even when the error wasn't simply the mis-setting of a parameter?

Furthermore, we have a picture in which strikingly different effects in child language are seen to be due to the same cause (e.g. the subsumption of the Korean word order error to the UCC). We don't have to search for a different cause for every different kind of child error, a particularly unhappy situation for a field that aspires to be a science. The field of child language begins to take on at least the hope that it might aspire to the theoretical, empirical and methodological standards of the more traditional "hard" sciences.

Is the OI Stage Due to Learning? No

The big question is, why the OI stage? Suppose we take the OI stage to be accounted for by the UCC from the last section. Why does it hold of children? What causes the OI stage to end? Answer: The going away of the UCC. But what causes the UCC to go away?

The answer from traditional approaches to language acquisition, including traditional generative grammar approaches, is that the children *learn* to leave the OI stage, they learn that the UCC doesn't hold. In my view this traditional answer cannot be right. Learning is by definition a change in the cognitive system due to the informational content of experience; for example, children learn to spell *the* in English; they don't learn to have teeth (even if it turns out that teeth are strengthed by use; there is no informational content in using teeth). Learning is the picking up of information from the environment; it is influenced by many variables. Learning implies that the behavior under discussion follows the laws of learning, for example, that learning changes to match the input and that lots of clear input will result in learning that matches the input well. Emergence from the OI stage cannot be the result of learning. Here are 4 excellent reasons; there are more that we will discuss as we find further sources of evidence, in studies of the causes of learning, in behavioral genetics and in studies of impaired development.

(20) *Problems for the Hypothesis that Learning is the Cause of the Fading Away of the OI-stage*

a. The evidence available to the child for finiteness being required in main clauses is enormous, existing in all input sentences. Children hear thousands and thousands of finite sentences, and very few sentences with main clause non-finite verbs. There is a tendency to speak shortly to children (Newport, Gleitman and Gleitman 1977), so that children hear fewer subordinate, potentially infinitival clauses than adults. At any rate, all sentences have a finite verb in the main clause. It is difficult to see what kind of a learning mechanism could be so faulty that it took several years to learn that finiteness was required. This is especially so since there is excellent evidence that children know the inflectional morphemes, with their grammatical and semantic properties, in the OI stage. E.g. they know that *s* in English can be used only with third person singular present tense verbs. So they have easily learned the properties of *s*. Except for one -- that it is obligatory rather than optional.

b. (i) If it is a question of learning, why should children start out mostly with forms that are not the most common forms in the input? Consider for example, the

83% OI rate in 1;6-2;0 Dutch children in (10). Since children hear so many finite verbs, if a learning mechanism is responsible for emergence from the OI stage, how did children ever get *into* the stage? Why don't they over-use the *finite* morphemes, which are used so often? We know that children essentially never substitute a finite morpheme where an infinitival morpheme is required. (See, for example, Guasti (1994) for Italian.) Yet this is what we would expect if it were a question of learning which morpheme goes where. Remember that in most of the languages studied, the equivalent of infinitival *to* in English isn't used -- rather simply the infinitival verb is used, e.g. *kopen/buy (non-finite)* in Dutch (6). So the infinitival verb follows the direct object in (6). But finite verbs *also* follow the direct objects in embedded clauses -- they too occur at the end of sentences. So why don't children make the "learning" error of deciding that finite verbs can substitute for non-finite verbs? They don't. There just doesn't seem to be a learning mechanism with the properties that will capture the empirical facts.

(ii) The problem is even more acute in languages like Danish, Norwegian or Swedish. These are languages without surface agreement -- there is only one form for present tense, and it occurs in every present tense sentence, doesn't vary with the features of the subject. So this one "present tense" form is extremely frequent. Yet in these languages there is a very high rate of OI's in young children. Since the same present tense form is so frequent, why isn't it substituted for the infinitival form rather than the other way around? Again, what learning mechanism could possibly have the required empirical properties?

c. We have seen that basic sentence and inflectional parameters are learned extremely early and extremely well, with almost no observable error (VEPS). That is, parameter learning for these parameters is over, completed successfully by the time the child enters the 2 word stage, around 18 months of age. There is evidence even in some children learning of basic sentence parameters (such as V2) is successfully completed at a somewhat younger age than 18 months. Given that children's learning abilities are so outstanding that they have learned basic parameter values perfectly at such a young age, what is it about their learning mechanisms that is so poor and leaky that the obligatoriness of finiteness is only mastered a couple of years later? It is simply very difficult to put together the exquisite early learning of parameter values with the late learning of obligatory finiteness, if only one learning mechanism is to account for both properties.

Putting these (and many other arguments, we will go over some of these later) together, we see the great value in studying parameter-setting empirically in children. Namely, parameters are language-specific, their values vary depending on the language. Thus there is unanimous agreement among nativists and empiricists (even behaviorists) that the parameter values (or whatever accounts for this variation) have to be learned from experience; there is no question of that. I believe that there is excellent evidence(both theoretical (e.g. learnability arguments, see Wexler and Culicover 1980) and empirical) that many principles are genetically programmed. But empiricists deny this claim; they think that principles (to the extent they believe that principles exist) are learned. So that at the very least we have to say that the evidence if arguable, if for no other reason than that it takes an argument to claim that a principle is genetically programmed. But the claim that parameters are learned is incontrovertible.

So parameters are a perfect testing ground in which to study learning -- we know that they *must* be learned. If one wants to study learning, parameters (or other aspects of language where it is known that there is variation, for example varying properties of the lexicon) are the place to study them.

When we carry out this study, in OI analyses, we discover that children are brilliant, precocious learners. It was no surprise to anybody who studies the OI stage to see the result of Saffran, Aslin and Newport (1996) that showed that 8-month old children could learn some distributional properties of stimuli. What other than the ability to learn from such kind of evidence could underlie the ability of children before age 18 months (as measured by production data) to set their parameters correctly? It would be surprising if the ability emerged suddenly at, say, 15 months, resulting in correct parameter-settings in production at 18 months. (There are studies in some languages showing some word order patterns are produced correctly at 15 months). Children have to be able to attend to varying order of words and morphemes and perform calculations, including learning calculations, on these. It is good to have confirming experimental evidence at a somewhat younger age, it makes the world consistent. But if somebody carried out an experiment which showed that children *couldn't* learn distributional information at an age somewhat before 18 months, we would have to conclude that either the experiment didn't appropriately tap their learning ability, or the materials presented were too far from a language-like situation. For the evidence from the production data that children are excellent learners of this kind of information, at least in a language-like setting, is vast and overwhelming -- there is no way that this evidence could be consistent with a lack of learning ability.

d. If the OI stage is the result of a general human learning mechanism, we would expect that the OI stage would show up in second language learning by adults. It would simply be the result of applying a learning mechanism to input data. But in fact the OI stage does *not* show up in adults. The growing literature on this topic is relatively recent, but the evidence is already quite good. See, for example, Haznedar and Schwartz (1997), Prevost and White (1997) and Ionin and Wexler (in press). Adult L-2 learners *do* use root infinitives sometimes, but they have very different properties from OI's. For example, they often appear in second position in V2 languages -- something which never helps to OI's (Prevost and White). 5 to 10 year-old L-1 Russian speakers often consider finite forms of *be* to be a kind of tense marking, using *be* together with a stem form, *he is go* (Ionin and Wexler), something which children in the OI-stage almost never do (Rice and Wexler 1996). Haznedar and Schwartz show that even a young child (L1 Turkish) learning English continues giving lots of what appear to be OI's, but doesn't use null-subjects along with them, contrary to the behavior of children in the OI-stage. Ionin and Wexler replicate this result with their 5 to 10 year old L-1 Russian learners of English. I don't have the space to discuss this literature in any detail, but it seems that the best hypothesis is that adult L-2 learners have much more difficulty than young child L-1 learners in learning the exact properties of inflections (Prevost and White's hypothesis that adult L-2 learning has trouble with learning surface forms). Ionin and Wexler's review of the literature concludes that there is no OI stage in adult L2 learning. Adults *do* show some of the kind of error-filled, slow acquisition of morphemes and their properties that learning theories would expect. So at many points child L-1 learning and adult L-2 learning diverge -- the OI stage is not replicated in adults.

Is it Genetically-Guided Maturation? Yes

Fortunately, there are two answers available in science for what causes immature forms to grow into mature forms. Although learning plays a role in some instances, genetically-guided maturation is even more basic, and presumably more common. So the obvious hypothesis to make about the withering away of the OI stage, of the UCC, is that it matures away, under genetic guidance. In other words, the genetic system determines that at birth (or whenever the language system comes on-line) the UCC is in place and the genetic system also insures that the UCC dies out over time. The maturing away of the UCC is a matter of genetically-timed development, as are so many other aspects of development in both human and non-human biology.

Borer and Wexler (1987, 1992) made the classic arguments for maturation of the linguistic capacity in the generative tradition, and since then there has been a lively debate on the topic, a debate that I do not have the space to review. So here, let's just consider what evidence exists for the proposition that it is genetics that guides the withering away of the UCC and thus of the OI stage.

First, all the problems raised for the learning hypothesis in the previous section are easily dealt with by the hypothesis that the development is genetically-guided. Yes, children *are* excellent learners, as seen in their excellent abilities at learning the properties of inflectional morphemes like *s*. Children use their learning abilities to learn the features of *s* perfectly and early. But the UCC affects the child's ability to mark every root verb as finite. Genetic inheritance causes the UCC to be part of the young child's computational system of language (or to constrain it in some way), until it withers away, again under genetic guidance. So OI's can persist even though learning of features of morphemes (not constrained by the UCC) is finely-tuned. This solves problem (20a).

Problem (20b) asks why the child starts out with such a large proportion of OI's. All we have to assume is that the genetic system specifies that the UCC constrains the very young child's computational system of language and that it dies away over time, under genetic guidance. At the youngest age, the child is most susceptible to the UCC, and we see large OI rates -- the input didn't cause the OI rates, which are orthogonal to the input. This is exactly what is expected from genetically-guided systems. Forcing hard food into a child's mouth will not cause it to grow teeth, nor will saying lots of finite forms to the child at a very young age force the child to leave the OI stage. In general, maturational systems play out over time, in a graded, not usually discontinuous manner. Teeth grow, they get bigger. Similarly the effects of the UCC die away over time, so that OI rates will gradually diminish. See Lenneberg (1967) for examples of maturational curves in biology.

Problem (20c) is likewise no problem under the current view. Children set parameters correctly because their learning systems are so good, but this learning system won't solve the problem for them that the UCC in their brains (via genetically-based heredity) calculates that a sentence which needs double checking is ungrammatical, and that they therefore have to omit AGR or TNS, producing an OI. Infants are capable of learning much; they can't "learn" to grow teeth before their biology requires/allows it.

Problem (20d) also vanishes under the view that the UCC is a developmental constraint. Since adults (or older children) aren't subject to the UCC, second language acquisition by these older children or adults won't result in the properties of the OI stage. Whatever errors exist in learning a second language at an age past the OI age-range will be due to other factors, for example, the difficulty in learning language-particular material that adults show, a difficulty that very young children do not have.

Further Evidence that the OI Stage Dies Away under Maturational Guidance

So far we have given a number of empirical arguments from phenomena concerning facts of normal language development that show that the OI stage (the existence of the UCC) dies away under genetically-guided maturation. There is evidence from a wide variety of additional sources that shows that the development must essentially be genetically-guided maturation, and not a process of learning from experience. Again, it is important to point out carefully that the child *does* do a good deal of learning from experience, we have discussed some of the most striking evidence that exists that shows how good the child is at this. But development out of the OI stage is too slow, too delayed, too at odds with the input, to be an event of learning. The phenomenology of the OI stage is so striking when set alongside the background of the phenomenology of parameter-setting (learning) that it calls out for a different explanation. The "empirical footprints" of learning and maturation are fundamentally different (Babyonyshev et al 2001). If we are to pay attention to the empirical world, here we might notice that that world is standing up and waving at us.

I would like to discuss some additional arguments for maturation because they bring in a wide array of alternate methodologies and fields, and help us to integrate broadly across different empirical approaches to a major problem. At the same time, the last piece of evidence concerns Specific Language Impairment (SLI), so that we can even integrate impaired development into the picture, in an important way, and show how its properties flow from and contribute to our knowledge of normal development.

3 Additional Empirical Arguments that the UCC (OI stage) is Genetically (Maturationally, Developmentally), Guided:

(21) Variables that affect learning: The usual variables that affect learning of learned material (including learned material in language) do not affect the development of the UCC.

(22) Behavioral Genetics: The UCC develops more similarly in identical (monozygotic) twins than in fraternal (dizygotic) twins

(23) Specific Language Impairment: The UCC's withering away is greatly delayed in SLI, perhaps it never goes away. Moreover, children with SLI are excellent learners of material in language that needs to be learned.

Let's start with (21). Learning is influenced by many variables, as psychologists have shown for more than 100 years. Many of these variables are related to input and its properties. For example, richness of input leads to faster learning. If growth out of the OI/UCC stage is due to learning, we expect this growth to be influenced by the same variables that affect learning in general.

We have already argued conceptually in (20a) that it doesn't make sense to think that growth of finiteness is affected by richness of input. Nevertheless we can ask the question anyway. Perhaps we've missed the relevant property that makes input "rich." Perhaps there is some mysterious property of the input that doesn't always exist, and the child is waiting for this mysterious property to appear.

But we can adopt the strategy of finding out what variables affect learning in other domains of language, and see if those variables affect the learning of obligatory finiteness. If they don't, then growth out of the OI/UCC stage is not caused by learning, by any psychologist's definition of learning. Learning has to obey the laws of learning; if it doesn't it's not learning. For example, the growth of teeth is not a case of learning; this growth isn't affected by experience the way learning theory expects. So if we are to approach the question of learning in an objective, scientific fashion, we have to ask: do the variables that affect learning also affect the learning of obligatory finiteness?

Rice, Wexler and Hershberger (1998) carried out just this study in English. They had Rice and Wexler's sample of approximately 60 children (40 normal, 20 SLI) who had been studied longitudinally for several years, the normal children from 3;0 to 7:0, the children with SLI from 5;0 to 8;0. The question they asked was, what variables affected the growth of the obligatory nature of finiteness? It was straight-forward to quantify this variable; it's the percentage of finite forms used in obligatory contexts over a range of contexts, all of which are predicted to be sometimes non-finite in the OI stage. For example, omission of third person singular *s*, omission of *be* forms. Their results held for both normal children and children with SLI, so we won't separate those out here, but will return to the children with SLI when we discuss the nature of SLI.

To decide what variables to study as potential causes of the growth of finiteness, Rice, Wexler and Hershberger decided to test the variables that had been shown to strongly affect the growth of vocabulary size, to be predictive of the growth of vocabulary size. These variables had been taken to be important variables in causing learning to take place.

One variable was the amount of formal education that the mother of the child had. This *mother's education* variable had been shown to be quite predictive of growth of vocabulary in previous research (Huttenlocher 1991). And it makes a lot of sense. After all, vocabulary growth takes place in an item by item manner; it is normally thought to be influenced by number of presentations of the item, by the contexts in which it is presented, by the drawing of attention to objects and events, by richness of input in various ways. And the amount to which a parent does all these things is thought to be influenced by her degree of education, not categorically of course, but statistically, over the population. Vocabulary growth needs input, there can't be any question of that, and each item needs input. You can't learn a word you haven't heard or seen. So there's no question that growth of vocabulary is influenced by learning, at least a significant part of it is learning, and this is constant, since each item must be learned[13]. So this makes sense. *mother's education* was chosen because of its significant effect in the vocabulary studies; it was the most significant environmental variable found in those studies.

A second variable that has been shown to have a lot of effect on rate of vocabulary group in studies of growth of vocabulary the IQ of the child, *child's IQ*. This makes sense because IQ is considered to be related to general ability to learn. Since vocabulary growth has a large component that has to be learned and each item has to be learned, we would expect *child's IQ* to have be predictive of rate of vocabulary growth, and it is.

Rice, Wexler and Hershberger do hierarchical linear modeling -- an analysis of the same type as done in the studies of vocabulary growth -- to see what the effect of these variables are. The results are that -- in strong contrast to the results on vocabulary growth done using the same methodology --neither *mother's education* nor *child's IQ* were significantly predictive of the growth of the rate of finiteness. In fact, these two variables (*mother's education* and *child's IQ*) together with 3 other variables, including whether the child was in the normal or SLI group,

---

[13]There is very good reason to believe that much about the lexicon is part of UG and is genetically programmed (Jerry Fodor makes the extreme argument that everything about the lexicon is innate except phonetic spell-out). But no matter how much of the structure of the lexicon is innate, the phonetic spell-out plus the choice of which items *are* spelled out in the lexicon has to be learned item by item (short of productive rules, in the lexicon, what are sometimes called *lexical redundancy* rules).

together accounted for only .3% of a reduction in variance in the growth of finiteness, less than a third of one per cent!

This is a remarkably strong result, using just the kind of data and method that is needed to test the idea of whether the growth of finiteness follows the laws of learning, that is, is influenced by variables which influence learning. What the results tell us is that if you look at 2 children, with the same level of finiteness (in obligatory contexts), but one of whom has a higher IQ and a mother with more education than the other, you will know nothing at all about how to predict which of the 2 children has a faster rate of growth in finiteness! Finiteness grows independently of the mother's education or the child's IQ. The 2 children will likely grow at different rates because growth is not identical across children. But you won't know anything given the other variables about how to predict which will grow faster. The growth of finiteness simply contradicts the laws of learning; the growth is not learning.

The situation is comparable to the following. Suppose we have 2 children of the same age. Knowing the IQ of the child and mother's education level, will you be in a better prediction to predict which child's hair will turn gray earlier? Who knows, maybe there *is* an effect, but we wouldn't expect one intuitively; we wouldn't be surprised if mother's education and child's IQ did not influence when hair turned gray. The reason we wouldn't be surprised is that we don't believe that the hair's turning gray is a process of learning; it doesn't follow the laws of learning, so variables that affect learning shouldn't necessarily affect the hair's turning gray.

What variables *do* affect growth of finiteness? Simply, time. A linear function of the time that has passed reduces the variance in finiteness rate by 72% and adding in a quadratic function reduces the variance more than 87% That is, you know almost everything there is to know about a child's finiteness rate if you know at what level she is when you measure her the first time and how much time has passed since then. If you have 2 children with the same finiteness rate, and you measure their finiteness a year later, they will be very close in finiteness at the later measure; there is very little random fluctuation in growth, given the 87% reduction due to time. (If this number were 100%, then any 2 children who have the same rate at time t_1 would have to have the same rate at a later time t_2 -- there would be no statistical flux at all. So the 87% figure is huge.)

Of course the fact that the passage of time is the major factor(almost a complete factor) in growth of finiteness, and other variables are not factors at all, is exactly what is expected on a maturational model. As the passage of time occurs, and the child ages, the genetic system carried out its functions. The UCC dies away as time goes by, uninfluenced by the variables that affect learning, simply influenced by the passage of time, the effects of which arise from the genetic system[14].

What is particularly beautiful, almost surprising even to a theorist who believes that the principles of language *grow* rather than are learned (as Chomsky has often written) is the extent to which the empirical data, gathered via traditional quantitative psychological studies of longitudinal data, confirm the essential *growth* character of the demise of the OI/UCC stage. This looks like science; it looks like biology. It looks the way Eric Lenneberg's classic *Biological Foundations of Language* (1967) expected language development to look, although the developmental evidence didn't exist at the time. Perhaps we should think of it as Lenneberg's dream.

Behavioral Genetics

Turning to the behavioral genetic data (22), we can ask the same question as we've just discussed, but turned on its head. In studying the question of which variables affect growth of finiteness (21), we were asking (this is simplified), if 2 children start out with the same rate of finiteness, what predicts differential rates of growth? Behavioral genetics asks, if 2 children are identical in genetic system to such and such an extent, how much does this genetic identity predict a growth similarity compared to the growth similarity of 2 children who are less identical?

Ganger and Wexler (1997) use the standard behavioral genetic method of studying a group of identical (monozygotic) and fraternal (dizygotic) twins. Essentially MZ twins share 100% of their genes and DZ twins share 50% of their genes, over a population. Ganger and Wexler studied the growth of finiteness in sets of these

[14]In principle it is possible that orthogonal factors are responsible for the demise of the OI stage. For example, perhaps the OI stage is due in some way to an immature pragmatic system. As this system develops the OI stage goes away. Although this is conceptually possible, there are severe empirical hurdles for such a proposal. For example, why don't children developing null-subject languages like Italian show the same pragmatic deficit, thereby producing OI's in their language? Even if the empirical challenge can be met by some refined theory, we will still have to ask the learning/development question of the pragmatic system. Can it be learned? Or is itself subject to maturation and to developmental constraints? At the moment I know of no proposals that solve these problems, or a sufficient body of empirical analysis (e.g. what variables influence the development of pragmatics?), but the question is ultimately an empirical one.

twins. To the extent that genetic factors affect the growth of finiteness, we expect that the MZ twins will be more similar in their development than the DZ twins.

The reason that the twin methodology is used in behavioral genetics is that it is assumed that both members of a pair of twins will grow up in a fairly similar environment, so that effects of the environment are controlled for. More essentially, it is assumed that whatever environmental differences there are between identical twins will not be exaggerated for fraternal twins. The methodology rests on that assumption. After all, siblings share the same proportion of genes (50% over a population). But the crucial assumption/hope is that fraternal twins, being twins of the same age, living in the same family environment at the same time, are treated as similarly as identical twins, who also are twins of the same age, living in the same family environment at the same time.

There are certainly cases where it is reasonable to question that assumption. To take an extreme case, suppose we discover that identical twins tend to dress more identically than fraternal twins. We wouldn't conclude that how one dresses has a genetic component, because it seems reasonable to guess that parents of identical twins might try to exaggerate their identicalness by dressing them alike, so that choice of dress is influenced by an environmental variable, parental training.

The argument of behavioral genetics rests on the assumption that we are studying a different kind of case, one in which the dependent variables that we are testing are such that the parents of identical twins are not any more likely to treat them similarly than are the parents of fraternal twins. Thus for vocabulary growth, say, the assumption would be that parents of identical twins are not more likely to give their twins a similar environment that is related to training on vocabulary than are the parents of fraternal twins. One can question this assumption and critics of behavioral genetics have often questioned the assumption, reasonably in many cases, in my opinion.

Vocabulary growth is a good example of how it might be possible for parents to affect the similarity of twins. It's conceivable, at least, that parents try to introduce words to each of two identical twins in a similar manner, and they have a much smaller tendency to do this for fraternal twins. I don't know whether to believe this or not, but it is certainly conceivable.

So we should approach all behavioral genetic data and analysis with a reasonable degree of skepticism. I would suggest, however, that if there is *any* cognitive or linguistic area where the crucial assumption if warranted, it might be the growth of finiteness. I have already given conceptual arguments that training differences shouldn't be relevant to growth of finiteness -- there are so many exemplars given to any child in a reasonably normal environment. What would the child *do* with more examples? The OI's don't come from what parents do, so it isn't as if parents choose a rate of OI's they're going to use and parents of identical twins would use a similar rate of OI's in talking to their 2 twins, whereas parents of fraternal twins wouldn't. Moreover, we have already seen that there is good data that shows that the intuitively plausible environmental variables don't affect growth of finiteness. These variables include *mother's* education which is presumably a surrogate for the things that a mother actually does to affect the child's environment. So it looks as if environment in the standard sense doesn't have any effect on rate of growth of finiteness. Thus to the extent that one accepts the behavioral genetic methodology at all, the growth of finiteness is exactly the kind of variable that can be studied relatively worry-free that a fundamental assumption of the method is being violated.

Ganger and Wexler studied a set of MZ twins and a set of DZ twins and measured how closely the twins in a pair attained a criterion in the use of obligatory finiteness. The measure was the difference of age of the twins when they reached the criterion. 0 weeks would mean that the twins reached the criterion at exactly the same time, and as the number grows, the more different the twins are in reaching criterion. The result turned out to be 13 weeks for the DZ twins, and 3 weeks for the MZ twins, In other words, the identical twins attained a criterion for a rate of finiteness on the average (over the set of identical twins) only 3 weeks apart; this number shot up to 13 weeks for the DZ twins.

Although preliminary, because it is the first behavioral genetic study of the growth of a property tightly bound up with early grammar, the result is quite promising. Ganger (1998) provides more evidence on this issue, using the same twin method. Much remains to be done, but to the extent that we have any evidence from behavioral genetics, we have evidence for the proposition that genetic variation affects rate of growth of finiteness. We can conclude that rate of growth of finiteness is affected by the genetic system. This is what would be expected on a maturational (growth) view of the development of grammar. Some children develop faster because their genetic systems develop somewhat faster.

It should go without saying that we have no reason to think that children whose genetic systems cause their rates of finiteness to grow faster than other children's are superior in any way, or that their linguistic systems are superior. The situation is just like with rate of growth of bodily organs. All normal children develop; the rate of growth varies a bit. There is no question of superiority. Moreover, unlike continuous variables like height, the use of obligatory finiteness rises to the same rate --100%-- for the approximately 95% (see a later section on SLI) of normal children. Unlike height, use of finiteness at maturity does not show a normal distribution. The phenomenology is more like that for having a heart, with all its parts. Short of pathology, people develop hearts. Some grow faster than others, but people get there.

Specific Language Impairment

We finally turn to the study of Specific Language Impairment (SLI). SLI by definition is an impairment that is specific to *language*; children are considered to be children with SLI if they have any kind of cognitive or auditory or speech deficit. There seems to be a good-sized group of such children, approximately 5% of the developing population according to a large epidemiological study (Tomblin 1996).

Many chapters in this volume review the literature on SLI. But I want to describe its central features, and to relate these features to question (23), the fact that the UCC remains active far longer in children with SLI than in normal children despite the fact that **children** with SLI are excellent learners of linguistic material. It turns out that SLI is an impairment that strongly supports the genetically-guided maturational basis of the growth out of the OI/UCC stage, so I will concentrate on those features that are relevant to these questions.

One of the focus points of this volume is the study of SLI and the connection of this study to linguistics. As I pointed out in the beginning of this chapter, to study impairment in some domain of language, we must have a good idea of normal development, its technical features, its structures and how they are attained, and what mechanisms drive this development. The study of the OI/UCC stage has all these features; I believe that it is understood better in technical detail, with the integration of a range of empirical material, than any other domain of early linguistic development. Moreover, I believe that we have more clear empirical information --much of which I have discussed -- about the mechanisms that drive the growth of language in this domain than in other domains[15]. I believe (with

[15]In general we have a better understanding of developmental properties of language that have been described within the Principles and Parameters approach than those which haven't. The idea of genetically-driven maturation of parts of the computational system of language was actually introduced into language acquisition studies with the results on passives and related structures (Borer and Wexler 1987, 1992, Babyonyshev et al 2001, Miyamoto et al 1999, Lee and Wexler (in press)). But in the case of the OI stage we have a great deal of added quantitative evidence about

linguistic theory and with researches in language acquisition itself) that much of linguistic growth, outside of the domain of experienced-based language variation, like parameters, is driven by genetically-based growth.

Thus when I decided with Mabel Rice to undertake a study of SLI, it seemed only natural to ask whether the **children** with SLI were in the OI stage for too long a period, and how much of their behavior could be accounted for by this very simple hypothesis. This is our *Extended Optional Infinitive Hypothesis* (also see Rice this volume), which says that children with SLI are just like normal children except that they go through the OI period for a much longer time than normal children, perhaps never really emerging from it. Given that the OI period is more accurately (on the current theory that I have discussed here) a period in which the UCC holds, we could call the stage the Extended Unique Checking Constraint (EUCC) period. The name doesn't matter, but the assumption does. The hypothesis is that whatever causes the OI stage is present in children with SLI for a much longer time, perhaps indefinitely.

Why was this natural? Because I had already decided that the best hypothesis about normal development was that the OI stage was the result of a genetically-driven maturational stage. Thus it was natural to believe that the genetically-driven event that caused the demise of the OI/UCC stage didn't take place or took place late in children with SLI. The mis-timing of genetic events is well-known enough to have a name in the genetics literature: heterochronology. So it was a natural enough biological possibility.

Of course, the naturalness of the idea didn't mean it was true. It was almost too much to hope for that such a simple idea could turn out to be true. Wouldn't it be more likely that SLI grammar was far more different from normal grammar than just in the processes that underlie the OI stage? This was a brute empirical question, and it received a very simple and clear answer in the work that I've done with Mabel Rice in English. The EOI *does* characterize SLI.

Crucially, in order to demonstrate that the EOI characterizes SLI, one has to show much more than that children with SLI produce too many OI's for their age. That result is necessary but not sufficient. Recall that the OI/UCC stage is characterized by a number of features. One of the central properties of the OI/UCC stage is that

---

variables that cause learning, behavioral genetics, impairment studies, detailed relations to second language acquisition, etc. Part of the reason for this is the simplicity of the phenomena; I fully expect that the same kind of evidence will be available for more complex cases as research proceeds.

parameters have been correctly set. Another of the properties is that major inflectional morphemes in the verbal system have been learned correctly together with their syntactic and semantic features. In other words, in the OI stage children show a particular deficit (for example, the production of non-finite verbs in many languages) together with a range of excellent competence in other aspects of the computational system of language. (Of course, NS/OI predicts that SLI in Italian-like languages will not show OI's. We will return to a discussion of what the theory predicts to be a marker for SLI in a language such as Italian). It is crucial to determine that children have this knowledge/competence alongside the specific deficit if one is to argue that children are in the OI stage.

Following this reasoning, Rice and I decided to study the EOI stage by both studying the phenomena which were predicted to show a deficit (finiteness marking on verbs) *and* the phenomena which were *not* predicted to show a deficit. For the latter we chose as the first piece of competence to look at the question of subject-verb agreement. Children in the OI/UCC stage get subject-verb agreement right, in the sense that if a child uses a finiteness morpheme, the subject almost always agrees with this morpheme.

This was first shown for German by Poeppel and Wexler (1993). For example, we showed both in our data and in other data in the literature (Clahsen1986) that when a child used third person singular *s* the probability was greater than .97 that the subject was third person singular. When the child used the morpheme on the verb for first person singular, the probability was similarly great that the subject was first person singular. The child knew the agreement morphemes and their features, so that the subject always agreed with the verb. This was an essential part of the OI stage.[16] The essential property is that the child has stored the verbal morpheme together with its correct (adult) features.

Similarly, Harris and Wexler (1996) shows that English-speaking children in the OI stage never used *s* with anything other than a third person singular subject. Very young children learn correctly the features that go with verbal suffixes.

Rice, Wexler and Cleave (1995), in the first empirical study of SLI in terms that took account of the OI stage, showed that two central properties of the OI stage held in children with SLI who were much older than the normal OI range.

---

[16]Given the ATOM, there can be more subtle predictions about a language. See Wexler, Schaeffer and Bol for a discussion of how ATOM might predict agreement errors in Dutch OI children, for example. But even on this latter analysis, the child has inserted into her lexicon the verbal agreement morpheme together with its correct features (person, number). Buy agreement or tense features may be omitted from the structure, producing the errors.

(24) English-speaking children with SLI at an older age than normal children:
a. Produce OI's
b. When the verb is finite, produce a subject that agrees with it, almost all the time.

To the best of my knowledge, these phenomena hadn't been known. In a rough manner, (24a) might have been thought to be known; after all, SLI was supposed to be having trouble with morphology, and leaving out verbal morphemes was one way that this happened. It wasn't thought of as lacking finiteness, nevertheless the phenomenon itself wasn't surprising.

But (24b) was not only not known (to the best of my knowledge) in the SLI literature, but it went against the received opinion that said children with SLI had trouble with morphology, and that they had a learning deficit concerning morphology. For if children with SLI really *did* have a learning deficit in morphology, we would expect them to produce agreement errors. Since they sometimes used finiteness/agreement morphemes (like *s*), to have a learning deficit would mean on any kind of computational model, that they had stored *s* with potentially incorrect features, that it were used at least sometime in a random manner so that the subject might not agree with it; the children might say **I goes* or **they goes*. But this is exactly what does *not* happen in SLI, as Rice, Wexler and Cleave showed. children with SLI were like normal children in this regard.

The EOI is quite different from the suggestion that children with SLI drop morphemes, to get shorter forms. That might work for some phenomena in English, but it doesn't work in other languages. Remember when we showed that young normal Dutch produce large numbers of OI's (see the table in (10))? These OI's are not shorter forms than the correct agreement forms, they just substitute a different suffix morpheme(*en*) for the finite morpheme.(8a) repeated below as (25) is an example, with *en* added to the stem *wass*.

(25) pappa schoenen wassen
daddy shoes wash-INF
'Daddy wash (non-finite) shoes'

In fact, for the first person singular, the agreement morpheme is Ø, the inaudible, phonetically zero morpheme. So when children use OI's instead of first person singular verbs (there are large numbers of these; see data in Wexler, Schaeffer and

Bol in press), they are complicating the verb, they are *adding* material to it in a surface sense. So there is no empirically reasonable notion of "surface shortening" in SLI or in normal children in general (thus no empirically adequate defense of the "Surface Hypothesis" of Leonard (1989)) or of the ideas on "morpheme omission" in Bishop (1997)). The notion of "shortening" or "omission" of surface material was a pure accident of over-concentration on the study of English, where the infinitival morpheme is phonetically zero. As soon as one expands the range of study out to even the closest related languages (e.g. the Germanic languages, the Romance languages) one sees that shortening is not empirically correct.

So the general idea of the EOI (and ultimately of the EUCC) is that the UCC has not been eliminated via genetically-driven maturation in children with SLI, despite the fact that they are at the age where it *is* eliminated in normal children. But other grammatical development is intact[17]. Thus we predict in general for children with SLI:

(26) Children with SLI:
a. Use OI's in languages where younger normal children do
b. Show the same patterns of grammatical knowledge as normal children

But it's probably easier to describe the logic of establishing the EOI/EUCC by considering a language that had the kinds of properties that the original OI languages had, with surface infinitival morphemes and with processes of parameter-set verb movement that allowed for strong predictions of morphology/word order correlations.

---

[17]Actually, there is one other strongly natural possibility. It is quite possible that children with SLI are delayed not only in the OI/UCC, but also in other areas where normal children are themselves maturationally delayed. That is, it is possible that SLI shows delay from normal children on grammatical property P if and only if P is itself a property that matures in normal children. Call this the "Hypothesis of Delay in All Maturational Properties." For examples, there is good evidence that A-chains mature over time (until around 5 years of age) (the *A-chain Deficit Hypothesis* of Borer and Wexler 1987, 1992, Babyonyshev et al 2001, Lee and Wexler in press, Miyamoto et al 1999 among many others); this is a very well-known area of maturational delay in the computational system of language. If the Hypothesis of Delay in All Maturational Properties is correct, then we would expect children with SLI to be seriously delayed from normal children in the representations of A-chains, for example, verbal passives, being able to give verbs a correct unaccusative analysis, etc. There is preliminary evidence in unpublished research that Mabel Rice and I are doing that there is not much serious delay in verbal passive of children with SLI (certainly they are not delayed compared to language(MLU)-matched controls, whereas the central results of Rice, Wexler and Cleave (1995), Rice and Wexler (1996) and many other papers is that children with SLI *are* delayed on finiteness rates relative to language(MLU)-matched controls). To the extent that English-speaking children with SLI are not delayed on verbal passive and similar structures, the strict EOI/EUCC is correct -- it is only UCC-implicated structures on which children with SLI are delayed. To the extent that children with SLI *are* delayed on verbal passive and similar structures, we will need a loosening of the EOI hypothesis to allow for delay on A-chains and similar. The logic of the 2 hypotheses is quite clear; they are both natural. Future research will decide which is more correct. At any rate, the fact that children with SLI are OI/UCC delayed is quite well-established.

Consider Dutch. In (10) we showed that normal children in Dutch go through an OI stage that is largely over in the 3;0-3;6 age-range; in that interval there are only 7% OI's. In the same paper, Wexler, Schaeffer and Bol studied 20 children with SLI. In the 6;00-8;02 year range, the children with SLI still had 15% OI's (50 of 334). The OI stage persisted much longer in the children with SLI. This is property (26a).

But especially striking is that correlation between verb second position and morphology. We showed this for normal children in (11). Here is the table from Wexler, Schaeffer and Bol for children with SLI:

(27) *Finiteness/position contingency Children with SLI*

all children with SLI	**V2**	**Vfinal**
finite	1071 (99.8%)	16 (5%)
non-finite	2 (0.2%)	335 (95%)

The data is remarkable because the children with SLI are so obviously excellent at the essential correlation. 99.8% of all V2 verbs are finite. But only 5% of final verbs are finite. This is beautifully precise, with very little having to be accounted for by performance or measurement error, at most 18 items out of 1,424 items (again, only non-ambiguous data cases were counted). Children with SLI are essentially perfect at the correlation, they're essentially just like normal children. This is exactly what (26b) predicts; it is an essential part of the OI stage!

This has to come as a surprise to any model of children with SLI which says that they're lacking grammar, or lacking the ability to learn surface morphemes (how could they get this correlation so perfectly right if there was something they hadn't "learned" about a surface morpheme?) Note that there is no question even of "omitting" morphemes. The non-finite forms have an *en* ending. The finite forms have a *t* ending in 2nd and third person singular, of which there are plenty (see detailed tables in Wexler, Schaeffer and Bol). Only the first person singular finite forms have a zero ending, and these are in *second* position since they're finite. So the OI's actually make some verbs *longer*, and moreover they get put in final, unmarked position, where non-finite verbs go.

So the Dutch children with SLI are clearly in the EOI/UCC stage. (They get agreement perfectly right; there is excellent evidence in the original paper that they have stored the agreement morphemes with the correct features).

Considering only the English and Dutch cases, we can already see that we now know much in technical detail about the nature of knowledge *and* non-knowledge in **children** with SLI. Moreover, we now know much about the effect of a *learning deficit* on SLI. With respect to the computational system of language there is *no* learning deficit! For Dutch children with SLI have set their parameters completely correctly. They get the V2/finiteness correlation perfectly; they behave completely correctly with respect to the *yes* setting of the V2 parameter, which Dutch exhibits. No SLI child has failed to learn that value of this parameter, and they hardly even show any noise on behaving with respect to the correct parameter value.

So children with SLI are *brilliant learners*, just as normal children are. They learn the language-particular properties that have to be mastered. They do not have a learning deficit.

Dutch and English-speaking children with SLI *are* delayed. There is a maturational delay in a property that is not a property that is learned, the property of obligatoriness of TENSE. That is, there is a delay in the demise of the UCC. Children with SLI at a much older stage are still governed by the UCC.

This Dutch and English data on SLI thus provide us themselves with a strong argument that the development out of the OI stage is genetically-driven maturation (23). For the children do not have a learning delay (parameters, agreement morphemes). When they *have* to learn, they learn, early, quickly, well. They pay attention; there is no attention deficit with respect to grammar; after all, they have to pay attention to learn parameter values.

How is it that we were able to draw such strong conclusions about the ability of SLI children to learn linguistic properties, in a field that has traditionally characterized SLI children as having a learning deficit with respect to language? Because we started with a clear idea of what was particular (parameters) and universal (principles) in language, and we ask the question, how do children perform on aspects of language that uncontroversially are learned -- parameters? We see that children with SLI learn parameters essentially perfectly; if there is a piece of language-particular information that normal children learn well, then so do children with SLI. As has happened so often, in every science, drawing fundamental distinctions (in this case between the definitely learned and the possibly/probably not learned) gave us a clear answer to a fundamental question. We now know that children with SLI do not have a learning deficit.

Clinical Markers for SLI: Cross-Linguistic Variation

It's crucial to have clinical markers for SLI, so that we can determine which children have SLI, both for scientific and practical reasons. Rice and Wexler have done extensive research arguing that rate of fineness is by far the most correct and sensitive clinical marker for SLI that has been proposed. There is virtually no overlap at a given relevant age between normal children and children with SLI on rate of over-all tensing. See the figure in (28) from the data in Rice and Wexler (1996).

(28) INSERT THE FIGURE HERE

The sensitivity and specificity of this grammatical marker for SLI argues for its usefulness; it is extremely rare in studies of cognitive abilities to have such a powerful cognitive marker. Of course, these results argue even more for the EOI nature of SLI.

What is intriguing, however, is that it follows from the underlying theory of the OI stage that the EOI stage will show extremely different surface properties in different languages. For example, suppose Italian children with SLI undergo the EOI stage in Italian. We have already shown that Italian children in the OI age-range do not produce root infinitives, and this follows form the interaction of the UCC with the parameters-settings of Italian (the null-subject parameter setting in particular). Since we have argued that children with SLI learn their parameters very well, and without a deficit, we expect that Italian-speaking children with SLI will have correctly set the null-subject parameter to *yes*. Suppose Italian-speaking children with SLI are subject to UCC at a much older age than normal children. Given their null-subject parameter setting, however, the UCC predicts that these children will *not* produce a significant number of OI/s. This is a quite startling prediction: English-speaking children with SLI produce huge numbers of OI's; tense appears to be a problem. But we are predicting that Italian-speaking children with SLI, because they have learned the null-subject nature of Italian, will not produce such. It is a real test of theory.

What *would* be predicted to mark SLI in Italian? Should SLI exist at all in Italian? The answer is that any non-adult utterance caused by the UCC should mark SLI at a fairly late age in Italian (or any other language). Wexler (in press) argued that one such error is the omission of object clitics. An object clitic is a pronoun whose thematic role is related to object position (after the verb), but which appears in pre-verbal position, a Clitic Phrase (ClP). Some element (the clitic itself, or in current theories more often an invisible noun phrase (*pro*)) starts out in object position and winds up in ClP. But since the clitic must be checked for case (ACC case or DATIVE case), the invisible noun phrase also has to pass through an intermediate position (known as AGR-Object on some accounts) which assigns ACC case. So on standard accounts *pro* moves and checks twice, to AGR-Object, and then to ClP. These movements can be thought of as checking the D-feature of the empty element, checking it twice, with AGR-Object and with INFL. Wexler argues that UCC prevents this from happening, often resulting in the omission of ClP and thus of the clitic. Informally, the double movement isn't allowed by the UCC but if both movements don't occur, there is something wrong with ClP, it doesn't have a NP with the right object features in local relation to it (*pro*). So ClP (and thus the clitic) must be omitted to obtain a good structure. Thus omission of Romance object clitics is predicted to be a consequence of UCC and omission of object clitics for an extended period of time is predicted to be a marker of SLI.

Here are the predictions re Italian SLI:

(29) a. NO OI's for main verbs
b. Nevertheless, omission of auxiliaries (see Wexler 1998 for the argument for normal children, which carries over to SLI)
c. Good agreement (because children with SLI learn well)
d. Major omission rates of object clitics

The fact that SLI seems to present so differently in different languages has made the whole problem seem intractable. But we now have fundamental *reasons* why there *should* be differences in SLI behavior in different languages, based upon a clear understanding of particular properties of grammar, variation among grammars, children's learning abilities, and children's maturational states. Taking *all* of these properties into account, with independent evidence for each one, gives us a clear picture. All that remains is to decide whether it's true. So, how about Italian SLI?

Bottari, Cipriani and Chilosi (1996) present a study of OI's in Italian children with SLI with some normal controls. Of 27 children with SLI with expressive-receptive deficits (thus matching the standard definition of SLI, for example those used in the Rice and Wexler studies), 20 of the children produce no OI's at all![18] This is already major information, as children with SLI in English and Dutch produce many OI's. Of the 7 children who do produce OI's, quantitative estimates are only available for 3 of them, and the percentage of OI's (with age of child in parenthesis) is 7.5% (6;2-6;11), 8.8% (8;7), 11.6% ((8;0). Although they are larger than the numbers for the 3 control children who are studied, they are extremely small by standards of the OI languages. Moreover, it is crucial to remember that 20 of the 27 produced no OI's at all. If we calculate 0% for the 20 participants with no OI's, and these numbers for the 3 participants who OI's whose weight is measured, we find a mean of 27.9/23 = 1.2% OI use per child! The authors write (p. 81):

...if RIs [= Root Infinitives, another name for OI's], produced by Italian children with SLI were to be accounted for in terms of [a hypothesis that the Italian OI's are accounted for by the same mechanism as non-null-subject language OI's] their frequency would have to parallel the frequency of RIs produced by children with SLI speaking English, French or German. This prediction is completely falsified by the English and German data....

The authors go on to argue that the few OI's that do exist in Italian children with SLI are something else, not the product of the OI stage. At any rate, we see a huge disparity in rate of OI's between Italian on the one hand and English or Dutch on the other. In Italian children with SLI there are almost no OI's; they have to be sought out. In non-null-subject languages, they are an obvious strong phenomenon.

The prediction of the UCC plus the hypothesis that children with SLI (like normal children) set their parameters correctly is strongly confirmed. Children with SLI behave strikingly differently in Italian than in English, and we expect exactly this difference.

Bottari, Cipriani and Chilosi go on to show that Italian children with SLI essentially get verbal agreement close to perfect, again as our hypotheses predict.

We already discussed why Wexler (in press) argued that UCC predicts that object clitics should be omitted during the OI (UCC) stage and that extensive clitic omission should be a marker of Italian SLI during this stage. Bottari et al (1998)

---

[18]Unfortunately, the authors do not tell us the ages of these children, but the ages of the 7 children who do produce some OI's are 6;5-9;1.)

show that there is extensive clitic omission by Italian children with SLI. The 11 children with SLI (mean age 6;3, range 4;2-10.7) omit clitics at a mean rate of 41.1%, whereas the 2 much younger normal controls omit many fewer clitics: 10.l% at age 32-34 months for Raffaello and 20.8% at ages 27-29 months for Martina). Basically (see Wexler in press for a review of the empirical evidence across a number of languages) the clitic omission stage is pretty much over in the 2's for normal children, but it is still huge for children with SLI of mean age 6.3.

As expected, the phenomenon of extensive object clitic omission in SLI is also characteristic of French SLI (see Jakubowicz, Nash Rigaut and Gerard 1998)).

As we have discussed, the UCC *does* predict that auxiliaries will be omitted in the OI stage, even though infinitival main verbs won' t be produced. Thus we would expect Italian SLI to show a large amount of auxiliary omission. This is confirmed for children with SLI in Bottari, Cipriani, Chilosi and Pfanner 1998). The children (mean age 6;3) omit auxiliaries at a 67% mean rate, strongly confirming the prediction. Compare this with the 1.2% OI rate discussed earlier). The two (much younger) normal younger in Bottari et al (1998) also omit auxiliaries, but fewer than the children with SLI, as expected. The predictions of the EOI/EUCC model are strongly confirmed. Italian children have their own pattern of deficit, which follows from the UCC restrictions, principles of grammar, and the parameter values for Italian that they have learned so well.

In general, we will expect different SLI behavior in different languages, and we have to be on the lookout for the phenomena that might be predicted by the theory[19]. Thus we expect the clinical marker of SLI in Italian to look quite different from the one for English, or for Dutch. The clinical marker should follow from the theory and the nature of each language. It is no surprise to our theory that children with SLI present so differently (on the surface) in different languages.

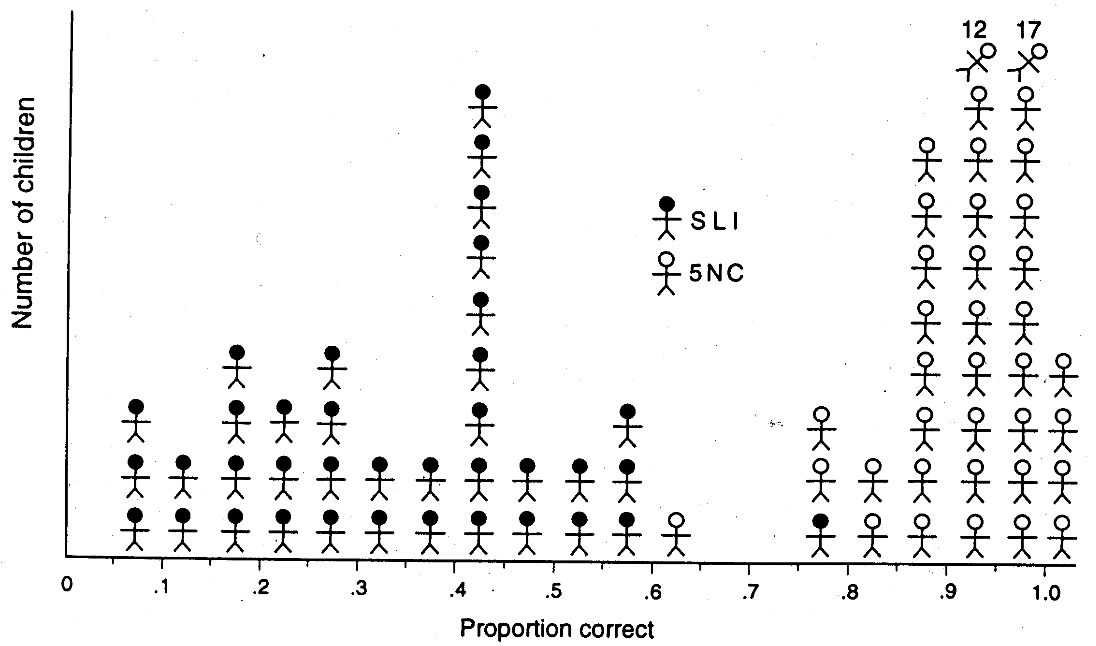
[19] For example, in Danish (and French) the UCC predicts other interesting patterns (for example the use of null-subjects with finite verbs in these non-null-subject languages (Wexler 2000)), which are well-confirmed (Hamann and Plunkett 1997). So this might play a role in the clinical marker for Danish SLI. In Korean, as we have discussed, Baek and Wexler (1999) argued that the word order error between *an/not* and the direct object was the result of the UCC. So although Korean doesn' t even have an infinitive (and we wouldn' t necessarily expect OI' s, since Korean might be a null-subject language), we might expect to see *an* mis-placement errors as a strong feature of Korean SLI. I don' t know whether these predictions are true. If they are not, it would argue that the UCC analysis given of these phenomena is wrong; we see how impairment data can affect our analysis of normal language development.

Underneath, they suffer from a common impairment, the extra restrictions on their computational systems caused by the UCC. On the surface, they look different.

This is no more surprising than that different molecules have different properties, although they all obey chemical law. The structure of a molecule will lead to different behavior, consistent with universal physical principle. I think it fair to say that the structure of the theory that we have discussed and its precise empirical verification make the science look more and more like chemistry, rather than like traditional psychology or the other social sciences. It is good to know that we can understand with such predictive precision what appeared to be possibly intractable problems. And -- best of all -- the answers aren't just some kind of statistical agglomeration coming out of a simulation which allows no insight. Rather, the empirical answers, combined with the theoretical analysis, allows us to hope to be able to understand -- perhaps for the first time -- the exact role of learning and the exact role of genetics and heredity in development, including SLI development.

Genetics and SLI

There is evidence that SLI has a strong heretibility component (Rice, Haney and Wexler 1998). We are currently engaged in a search for the genetic locus of SLI. If we find such it might help with the extremely difficult question, one that to the best of my knowledge no discernible progress has been made on, the question of the neuroscience of SLI, what happens in the brains of children with SLI. Perhaps if can learn what genes are involved with SLI, we might be able to figure out what proteins these genes code for and then try to understand what happens structurally. At the moment this sounds almost like science fiction, but who knows when the right breakthrough will be made. In my opinion, it could happen. If it does happen, the kind of detailed work, clarifying every aspect of what SLI and normal children are capable of, distinguishing development from learning, comparing languages, etc., will be of the utmost importance. We can't discover the biological basis of SLI until we understand its computational basis. The fact that SLI is a genetic event (or lack of one, the withering away of the UCC) is quite consistent with the observed genetic influence on the likelihood of having SLI. The world is consistent so far, but I have no doubt that there are all sorts of scholars working away to make it (temporarily) non-consistent, and thus to push us in new directions.



Number of children
Proportion correct
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
12
17
SLI
5NC

References

Babyonyshev M., Ganger J., Pesetsky D., & Wexler K., (2001). The Maturation of Grammatical Principles: Evidence from Russian Unaccusatives. Linguistic Inquiry, 32, 1.

Bertolo, S., Broihier K., Gibson, E., & Wexler, K. (1997). Cue-based learners in parametric language systems: application of general results to a recently proposed learning algorithm based on unambiguous "superparsing." Manuscript. Dept of Brain and Cognitive Sciences, MIT.

Bishop, D.V.M. (1997). Uncommon understanding: Development and disorders of language in comprehension in children. East Sussex, UK: Psychology Press.

Borer, H., & Wexler K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.). Parameter Setting (pp.123-172). Dordrecht: Reidel.

Borer, H., & Wexler K. (1992). Bi-unique relations and the maturation of grammatical principles. NLLT, 10, 147-189.

Bottari, P., Cipriani, P., & A.M. Chilosi. (1996). Root Infinitives in Italian SLI Children In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (Eds.). BUCLD 20 (pp. 75-86). Somerville MA: Cascadilla Press.

Bottari, P., Cipriani P., Chilosi A.M., & Pfanner, L. (1998). The Determiner System in a Group of Italian Children with SLI. Language Acquisition, 7(2-4), 285-315.

Bromberg, H. & Wexler, K. (1995). Null subjects in child Wh-questions. In Shutze, C., J. Ganger and K. Broihier (Eds.). Papers on Language Processing and Acquisition. MIT Working Papers in Linguistics, 26. Cambridge, MA: MIT Press.

Brown R., (1973). A First Language. Cambridge, MA: Harvard Univ. Press.

Chomsky, N. (1995). The Minimalist Program, Cambridge, MA: MIT Press.

Clahsen, H (1986). Verb Inflections in German child language: Acquisition of agreement markings and the functions they encode. Linguistics, 26, 79-121.

Clark, R. & Roberts, I. (!993). A computational Model of Language Learnability and Language Change. Linguistic inquiry 24 299-345.

Dresher, B.E. (1999). Charting the Learning Path: Cues to Parameter Setting Linguistic Inquiry 30: 1 27-67.

Fox, D, & Grodzinsky, Y. (1998). Children's passive: A view from the by-phrase. Linguistic Inquiry 29, 311-332.

Fodor, J.D. (1998). Unambiguous Triggers Linguistic Inquiry 29 1-36.

Ganger, J. (1998). Genes and Environment in Language Acquisition: A Study of Vocabulary and Syntactic Development in Twins. Doctoral dissertation. MIT.

Ganger, J., Wexler, K, & M. Soderstrom (1997). The Genetic Basis for the Development of Tense: A Preliminary Report on a Twin Study. In E. Hughes, M. Hughes & A. Greenhill (Eds.). BUCLID 21 (pp. 224-234). Somerville MA: Cascadilla Press.

Gibson, E . & Wexler K. (1994). Triggers. Linguistic Inquiry, 25,3, 407-454.

Guasti, M (1994). Verb syntax in Italian child grammar: Finite and nonfinite verbs. Language Acquisition 3, 1-40.

Hamburger, H. & Wexler K. (1973). Identifiability of a class of transformational grammars. In K.H.H. Hintikka, E. Moravcsik, & P. Suppers. (Eds.). Approaches to Natural Language Proceedings of the 1970 Stanford Workshop on Grammar and Semantics. Dordrecht: Reidel.

Hamann, C,. & Plunkett, K. (1997). Subject Omission in Child Danish. In E. Hughes M. Hughes & A. Greenhill (Eds.). BUCLID 21 (pp.220-231). Somerville MA: Cascadilla Press.

Harris, T., & Wexler, K. (1996). The Optional-Infinitive Stage in Child English: Evidence from Negation. In H. Clahsen (ed.). Generative Perspective on Language Acquisition, Language Acquisition and Language disorders 14. Philadelphia, PA: John Benjamins.

Haznedar, B. and Schwartz, B. (1997). Are there optional Infinitives in Child L2 Acquisition? In E. Hughes, M. Hughes & A. Greenhill (Eds.). BUCLID 21 (pp. 257-268). Sommerville, MA: Cascadilla Press.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early Vocabulary Growth: relation to language input and gender. Developmental Psychology. 27, 236-248.

Ionin, T. & Wexler, K. (2000). L1-Russian Children Learning English: Tense and Overgeneration of "Be". To appear in the Proceedings of the Second Language Research Forum, Madison, WI.

Jakubowicz, C., Nash L., Rigaut, C. & Gerard, C. L. (1998). Determiners and Clitic Pronouns in French-Speaking Children with SLI. Language Acquisition 7(2-4). 113-160.

Lee and Wexler (in press) Nominative Case Omission and Unaccusatives in Korean Acquisition. (have e-mailed HyoJinn)

Lenneberg, E. (1967). *Biological Foundations of Language*. New York: Wiley.

Loeb, D.F. & Leonard, L.B. (1991). Subject case marking and verb morphology in normally developing and specifically language-impared children. Journal of Speech and Hearing Research, 34, 340-346.

Leonard, L.B. (1989). Language learnability and specific language impairment in children. Applied Psycholinguistics, 10, 179-202.

MacWhinney, B., & Snow, C. (1985). The Child Language Data Exchange System. Journal of Child Language 12 271-296.

Maratsos, M., Fox, D. Becker, J A. & Chalkley, M.A. (1983). Semantic restrictions on children's early passive. Cognition 19, 167-191.

Miyamoto, E. T., Wexler, K. Aikawa, T. & Miagawa, S. (1999). Case Dropping and Unaccusatives in Japanese Acquisition. In A. Greenhill, H. Littlefield, & C. Tano (Eds.). BUCLD 23, 2 Sommerville, MA: Cascadilla Press.

Newport, E., Gleitman, H. & Gleitman, L. (1977). Mother, I'd rather do it myself. Some effects and non-effects of maternal speech style. In C.E. Snow & C. A. Ferguson. (Eds.). Talking to Children New York: Cambridge University Press.

Prevost, P. (1997). Truncation in Second Language Acquistion, Doctoral dissertation, McGill University.

Poeppel, D. & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. Language, 69:1, 1-33.

Rice, M. & Wexler, K. (1996). Toward Tense as a Clinical Marker of Specific Language Impairment in English-speaking Children. Journal or Speech and Hearing Research, 39, 1239-1257.

Rice. M., Wexler, K. & Cleave, P. (1995). Specific language impairment as a period of extended optional infinitive. Journal of Speech and Hearing Research 38, 850-863.

Rice. M., Wexler, K. & Hershberger, S. (1998). Tense Over Time: The Longitudinal Course of Tense Acquisition in Children with Specific Language Impairment. Journal of Speech, Language and Hearing Research 41, 1412-1431.

Rice, M., Haney, K.R. & Wexler, K. (1998) Family Histories of children with SLI Who Show Extended Optional Infinitives. Journal of Speech, Language and Hearing Research, 41, 419-432.

Saffran R., Aslin R.N. & Newport E.L. (1996). Statistical Learning by 8-Month-Old Infants. Science, 274, 1926-1928.

Sano, T. & Hyams, N. (1994). Agreement, finiteness and the development of null arguments. In Merce & Gonzalez (Eds.). Proceedings of NELS 24 (pp.543-558). University of Massachusetts Amherst: GLSA.

Schütze, C. T. & Wexler, K. (1996). Subject case licensing and English root infinitives. In A In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (Eds.). BUCLD 20 (pp.670-681). Somerville MA: Cascadilla Press.

Schütze, C. T. and K. Wexler (2000) An Elicitation Study of Young English Children's Knowledge of Tense: Semantic and Syntactic Properties of Optional Infinitives. In S.C. Howell, S. Fish, T. Keith-Lucas (Eds.). BUCLD 24, 2 (pp.669-683) Somerville, MA: Cascadilla Press.

Tomasello (2000) Do young children have adult syntactic competence? Cognition, 74, 209-253.

Tomblin, J. B. (1996). Genetic and Environmental Contributions to the Risk for Specific Language Impairment. In M. Rice (Ed.). Toward a Genetics of Language, (pp. 191-210). Mawah, NJ: Lawrence Erlbaum Associates.

Wexler, K. (1982). A principle theory for language acquisition. In E. Wanner and L. Gleitman (Eds.). Language Acquisition: The State of the Art. (pp. 288-315). Cambridge: Cambridge University Press.

Wexler, K. (1990) Optional infinitives, head movement and the economy of derivations in child grammar. Paper presented at the Annual Meeting of the Society of Cognitive Science, MIT, Cambridge, Mass.

Wexler, K. (1992). Optional infinitives, head movement and the economy of derivation in child grammar. Occasional paper #45. Center for Cognitive Science, MIT, Cambridge, Mass.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot and N. Hornstein (Eds.). Verb Movement. (pp. 305-350). Cambridge University Press: Cambridge, U.K.,

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. Lingua 106. 23-79.

Wexler, K. (2000). Three Problems in the Theory of the Optional Infinitive Stage: Stage/ Individual Predicates, Eventive Verbs and Finite Null-Subjects. In Billery & Lillehaugen. WCCFL 19 Proceedings (pp. 560-573) Somerville, MA: Cascadilla Press.

Wexler, K. & Culicover, P. (1980). Formal principles of language acquisition. Cambridge, MA: MIT Press.

Wexler, K & Hamburger H. (1973). On the insufficiency of surface data for the learning of transformational languages. In , K.H.H. Hintikka, E Moravcsik, & P. Suppes. (Eds.). Approaches to Natural Language Proceedings of the 1970 Standard Workshop on Grammar and Semantics. (pp.16-179). Dordrecht: Reidel.

Wijnen, F. (1998). The temporal interpretation of Dutch children's root infinitivals: The effect of eventivity. First Language 18:379-402.