From
Models to Modules:

Studies
in
Cognitive Science
from the
McGill Workshops

Edited by

Irwin Gopnik
Myrna Gopnik

McGill University

THEORETICAL ISSUES IN COGNITIVE SCIENCE
ZENON W. PYLYSHYN, Series Editor

ABLEX PUBLISHING CORPORATION
NORWOOD, NEW JERSEY

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data
Main entry under title.

From models to modules.

(Theoretical issues in cognitive science)
Bibliography: p.
Includes index.
1. Cognition. 2. Psycholinguistics. I. Gopnik, I. (Irwin) II. Gopnik, Myrna.
BF311.F75 1986 153 85-20952
ISBN 0-89391-355-3

Ablex Publishing Corporation
355 Chestnut Street
Norwood, New Jersey 07648

Contents

Preface

The 15 papers in this volume are the result of a series of Cognitive Science Workshops held at McGill University during 1982 and 1983. The workshops were organized by members of the McGill Cognitive Sciences Group under the chairmanship of Myrna Gopnik and Carl Frederiksen, and were funded by a grant from the Alfred P. Sloan Foundation.

There were seven Workshops, each organized around one of the following topics: language acquisition and development; text and text processing; computer chess; grammars, parsers, and language comprehension; scientific reasoning and problem solving; language and the brain; and semantics. There were about 40 presenters and commentators in all, and a large, varied, but overlapping, audience made up predominantly of linguists, psychologists, philosophers, computer scientists, and neurologists.

It would be foolishly utopian to expect an overall theoretical framework for the cognitive sciences to emerge from such an interchange. What does emerge, however, are a number of suggestions, tentative though they may be, for constructing isomorphisms between various forms and levels of cognition, and a few shrewd guesses about where isomorphisms are likely, or unlikely, to be found. These positive and hopeful results led the organizers of the Workshops to feel that it would be worthwhile to share a gleaning of the available papers with their fellow cognitivists.

The Editors wish to acknowledge the invaluable organizational assistance of Mike Dillinger and the secretarial assistance of Anna Dillinger. A special word of thanks must also go to the Sloan Foundation whose generosity made the Workshops possible and to Gordon A. Maclachlan, Vice-Principal Research and Dean of the Faculty of Graduate Studies, for his moral and financial support.

Chapter 1

The Acquisition of Grammar*

Brian MacWhinney
John Anderson
Carnegie-Mellon University

1. INTRODUCTION

In this paper, we give a brief sketch of a model that gives a reasonably explicit account of the shape of the fundamental principles underlying language acquisition. We believe that, unless solid bridges are built between child language research and cognitive science in general, there is a real danger that many of the advances made in the area of child language research will be lost. Roger Brown (1977) explains our worries:

> Developmental psycholinguistics has enjoyed an enormous growth in research popularity...which, strange to say, may come to nothing. There have been greater research enthusiasms than this in psychology: Clark Hull's principles of behavior, the study of the Authoritarian personality, and, of course, Dissonance theory. And in all these cases, very little advance in knowledge took place....A danger in great research activity which we have not yet surmounted, but which we may surmount, is that a large quantity of frequently conflicting theory and data can become cognitively ugly and so repellent as to be swiftly deserted, its issues unresolved.

Brown's warning should not go unheeded. In order to avoid the fate of becoming "ugly," child language research must become more firmly grounded on general principles deriving from cognitive psychology and linguistic analysis. There is continuing research aimed at grounding language acquisition theory on principles of generative grammar (Wexler & Culicover, 1980). Such work is typically well-grounded in linguistic theory, but often fails to pay attention to the overall architecture of human cognition. Our own approach is to work within the constraints of both linguistic theory and cogni-

* Thanks to Joseph Stemberger for providing us with examples of speech error types. Thanks to Pat Langley for providing criticisms on an earlier draft of this chapter.

tive theory. Unlike Chomsky (1980), we assume that human cognition is unitary and that language is not a special "organ" of mind. We believe that language is special in the sense that it, more than any other system, has utilized virtually every major aspect of the general cognitive system (excepting perhaps aspects of visual processing). Because language has commandeered so much of the mind, its structure is quite complex. Despite this complexity, the pervasive utilization of old cognitive structures by the linguistic function means that language processing is governed by basic principles of cognitive processing and that the acquisition of language can be explained in terms of general learning principles. The overall framework we will adopt is that of production system theory (Newell & Simon, 1972). The particular instantiation of production system theory we support is that proposed by Anderson (1983) as ACT*.

Our approach constitutes a merger of two major lines of research on the acquisition of language by the young child. The first is John Anderson's (1977, 1981, 1983) LAS model of the acquisition of the rules of syntax. The second is Brian MacWhinney's (1975, 1978, 1982, in press) dialectic model of the acquisition of lexical structure. These two lines of research have made remarkably similar assumptions about the nature of the learning process and the overall cognitive system, while focusing on largely separate aspects of the language acquisition. Both of us have found that neither of these problems can really be addressed independently of the other. Issues of segmentation and the control of grammatical morphemes have played an important role in Anderson's formulation of LAS, whereas problems in affix order and case are centrally involved in MacWhinney's attempts to account for the develoment of morphology. It is clear that, from the viewpoint of the theory of language development, the formulation of a model that deals adequately with both syntax and word formation would be a major advance.

The model developed by MacWhinney for lexical processing is a parallel interactive activation system. Anderson, in developing a model of syntactic processing, emphasized the use of goals to impose a sequential structure on the application of productions. However, we have discovered that these two models are really quite alike. MacWhinney also requires a sequential discipline to organize the sequencing of phonemes and to permit orderly access to the lexicon. Moreover, the pattern-matcher for ACT*'s production system is a parallel interactive system, with many properties in common with MacWhinney's. By making a few significant changes in the architecture of ACT*, a single system can account for both syntactic and lexical processing. This merged model will enable us to better understand the way in which lexical processing interfaces with syntactic processing.

We believe that computational simulation is the most appropriate format for the building of the needed bridges between child language research and cognitive science in general. A computational model is being utilized in this research not as a replica of the human being, but as a means of attaining

explicitness in psycholinguistic research. The authors are well-aware of the problems inherent in the artificial intelligence approach to cognition (Chandrasekaran & Reeker, 1974; Dresher & Hornstein, 1976). No one should misconstrue the ontological or phenomenological (Wittgenstein, 1953) status of computer simulations. In particular, no one should attempt to view the computer as a brain. Simulation on present-day machines cannot duplicate mental processes (Hinton & Anderson, 1981), it can only mimic them Anderson, 1978). However, even this mimetic function allows us to compare alternative models by measuring their relative goodnesses-of-fit to a set of data. In this sense, current computational models are best viewed as devices for the evaluation of theoretical claims.

Our own research program builds upon an increasingly active current of research in the formal representation of language acquisition. Much of this research has focused on topics in grammatical inference and induction defined rather strictly within automata theory (Biermann & Feldman, 1972; Blum & Blum, 1975; Feldman, 1972; Gold, 1967; Horning, 1969; Pao, 1969). We take this work as representing the basic logical framework against which models of language acquisition must eventually be represented. There have also been formal demonstrations of learnability (Anderson, 1976; Baker, 1977; Culicover & Wexler, 1977; Hamburger & Wexler, 1973, 1975; Pinker, 1982; Wexler & Culicover, 1980; Wexler, Culicover, & Hamburger, 1975; Wexler & Hamburger, 1973) articulated within the framework of particular models of language structure and processing. Of equal importance is the tradition that has focused on the nature of learning heuristics (Block, Moulton, & Robinson, 1975; Braine, 1971; Harris, 1977; Kelley, 1967; Kiss, 1973; Klein & Kuppin, 1970; Langley, 1982; Levelt, 1975; MacWhinney, 1978; Miller, 1967; Reeker, 1976; Siklossy, 1972), often with a particular eye toward simulating data from child language acquisition. These various currents and their results are fully reviewed in MacWhinney (1978), Pinker (1979), and McMaster, Sampson, and King (1976). Unfortunately, we do not have space here to detail our indebtedness to each of these authors.

The following five sections summarize our model of language use and acquisition. The five topics that organize the presentation are as follows:

1. *General Architecture:* Here we review the general architecture of the ACT* system with a special focus on the parallel interactive processes underlying pattern matching.
2. *Lexical Activation:* We show how the various modes of lexical processing (rote, analogy, combination) can be understood within the single framework of a theory of parallel activation.
3. *Syntactic Processing:* We argue that the ordering of elements in sentence production is controlled by a serial goal-based syntactic process.
4. *Monitoring:* We show how the system operates in four modes to detect errors in comprehension and production.

5. *Acquisition:* Finally, we explain the development of new abilities in both lexical and syntactic processing in terms of a set of nine modes of acquisition.

2. GENERAL ARCHITECTURE

In ACT* behavior is characterized as being under the control of production rules which are condition-action pairs. For purposes of exposition let us consider the following production rules:

P1: IF the goal is to communicate: {[+obj][+anim][+fur][+canine]}
THEN set as subgoals to:
1. generate /d/ in onset position
2. generate /aw/ in nucleus position
3. generate /g/ in coda position

P2: If the goal is to communicate: [+obj],[+anim],[+fur],[+feline]
THEN set as subgoals
1. generate /k/ in a onset position
2. generate /ae/ in nucleus position
3. generate /t/ in coda position

For short-hand we will denote these rules as:

P1: [+obj][+anim][+fur][+canine]---->/d/ /aw/ /g/

P2: [+obj][+anim][+fur][+feline]---->/k/ /ae/ /t/

The semantic features given above are somewhat fanciful, but our argument does not depend on particulars of semantic analysis. The critical observation is simply that there will be partial overlap among the features underlying *dog* and *cat*.

Let us suppose the child's goal is to communicate the features [+obj], [+fur], [+anim], and [+feline]. Then there would be a partial match to the conditions of the two productions above. In ACT*, these features would be activated, and activation would spread from them to the two productions. Each production would be activated to the degree its conditions were matched by active features. In this example, P2 would be more activated. We believe that there is an inhibitory competition among productions like these which overlap in their conditions. Because of this, P2 would inhibit P1 and the spekaer would product /k/-/ae/-/t/ rather than /d/-/aw/-/g/.

In the theory developed in Anderson (1983), the productions would wait until a clear winner was determined in the inhibitory competition. Then the winning production would fire, and its action would be executed. The evidence from speech errors (Garrett, 1980; Stemberger, 1982) indicates that this architectural principle is wrong. Rather, the actions of the productions should be activated to the degree their production is active. By allowing for relative degrees of activations of actions, we can account for blends such as /d/-/ae/-/t/, where the /d/ comes from *dog* and the /ae-t/ comes from *cat*. Thus, rather than having productions fire in an all-or-none manner as in Anderson (1983), we now allow that productions be allowed to fire to varying intensities. This introduction of relative degree of production activation is a significant augmentation to the production system architecture. Other standard assumptions that go along with the notion of varying levels of activation are: (a) that, at any time, any cognitive unit or element has a non-negative level of activation associated with it; (b) that each unit has a resting level or strength; (c) that elements that reach a certain threshold of activation become "working memory" elements and continue to provide activation for a certain time; and (d) that spread of activation occurs in parallel both top-down and bottom-up.

The following factors determine the activation of various production rules:

1. *Strength of rules:* The strength of rules reflects the frequency and recency of their successful firing. Stronger rules receive more activation. To illustrate this, note that common irregulars like *went* tend to resist overregularization more than less common irregulars like *mend* (MacWhinney, 1978).
2. *Specificity:* The matcher rewards rules for having features matched. Thus, in word recognition, *cat* is better than *at* as a match to /k/-/ae/-/t/, since it matches three segments and *at* only matches two. A special case of this is the superiority of portmanteau forms to their analytic counterparts. In French, both *du* and *de+le* compete for the masculine partitive. However, because *du* is more specific, it gets more activation. Without support from "specificity," the less frequent form *du* might never win out over *de+le*. But, because of specificity, it is generally the case that rote forms dominate over combinatorial forms (MacWhinney, 1978, 1982).
3. *Accuracy:* The matcher penalizes rules for having too many features, i.e., for having features that are not active in working memory. Thus, when matching to /k/-/ae/-/t/, the item *bat* will be penalized for the failure to match a /b/.
4. *Data refractoriness:* The pattern matcher attempts to assign each active element in working memory to a single rule. This means that, if a particular element matches more than one rule (e.g., [+fur] matches P1 and P2), there will be an inhibitory relationship set up among these elements. This is what prevents multiple competing rules from applying to the same goal. As McClelland and Rumelhart (1981) note, this inhibitory relationship makes "the rich get richer and the poor get poorer" in that good guesses are supported and poor guesses eliminated.

5. *Top-down support:* If there is considerable activation of a particular action element, that activation will also support the production patterns that led to it. As we will see, such support can lead to the phenomena of haplology, malapropism, and analogy.

3. LEXICAL ACTIVATION

In the 1978 version of the dialectic model, MacWhinney proposed that lexical items could be generated by either *rote, analogy,* or *combination.* We will refer to the general process of lexical generation as *spellout.* In addition to these three basic processes, the process of *selection* works when two mutually exclusive alternatives have been activated. In this section, we will see how rote, analogy, combination, and allomorphic selection can be put into the framework of an interactive activation system such as the one described above. In effect, these modes of activation fall out as consequences of the principles governing pattern matching in ACT*.

One very important aspect of our current approach is the way in which we understand the interface between serial and parallel processes. We are hypothesizing that lexical activation occurs in parallel, and that much of syntactic processing is controlled by a serial mechanism. To interface these processes, we allow syntax to activate a data structure that controls phonological generation. No serial control is enforced while activation is taking place within an arena. Once units have been activated to confidence level, the data is then read out into articulation.

A great deal of child language theory has been based on the analysis of lexical errors made by children. In this section we will refer to errors made by children and errors made by adults. Later we will discuss these various error types as important phenomena that any model of natural language acquisition and processing most be able to explain. Our examples of speech error types are all taken from Stemberger (1982) and, in many places, our analysis closely parallels that of Stemberger (1982).

3.1 Rote

Rote application occurs when all the output segments of a given form are activated by a single production. For example, when lexicalization converts the meanings underlying *dog* to /dawg/ it must do so by rote, since the morpheme *dog* is an unanalysable unit. Productions P1 and P2 from the previous section instantiated rote rules for *dog* and *cat.* MacWhinney (1978, 1982) has surveyed a variety of developmental phenomena that indicate that higher-level forms such as complex words and phrases are often used in a rote fashion. Examples of these phenemona include:

1. Children often use words in phrases before they use them independently. Thus, *I could easily do that* can occur before *could,* and *can't* can occur before *can.*

2. It often happens that children start to make overgeneralizations for forms that they earlier produced without error. Thus, at first, children say *went* and *jumped;* later they say **goed* and *jumped.* This suggests that *jumped* is being controlled first by rote and then by combination.
3. Children often use both words and longer phrases without superfluous markings. Thus, *say bye-bye* may mean "bye-bye", *hat on gone now* may mean "hat gone now", and "brush your teeth" may mean "toothbrush".
4. Failing to analyse phrases may also lead children to use pronouns incorrectly, as in *I carry you* instead of *you carry me.*
5. Children may use affixes and other operators in contradictory ways, as in *shoes on off.*
6. Sometimes, children may produce whole sentences far in advance of their usual productions, as in *No, Mommy, I don't want to go to bed* produced by MacWhinney's son Ross during the two-word-stage.
7. In both children and adults, frequent words are less subject to overregularization than infrequent words. Thus, *went* resists **goed* better than *thrust* resists **thrusted.*
8. Related items have an intonational/clitic integration. For example, *want to* and *let me* have become *wanna* and *lemme.*
9. The semantic class of an item may not be coded at first. This can lead to *rices* for "rice grains", or *many lions,* then **many lion,* and then *many lions* again.
10. Correct wholes may precede their analysed pieces. For example, *drop* may be used causatively before the appearance of *I made it fall* or *I felled it.*
11. At first, polysemy may not be detected in compounds. Thus, MacWhinney's son Ross confused his friend Matt with the doormat and the button of a shirt with the button on a television.

3.2 Analogy

In a system with interactive activation within a given arena, analogy can occur when one or more rhyming items obtain enough top-down support to activate output segments that would not otherwise be activated. For example, the production of the /z/ of /p/-/l/-/u/-/g/-/z/ may arise from top-down support for /r/-/u/-/g/-/z/, /b/-/u/-/g/-/z/, and so on. These forms also receive activation from [+pl], giving them enough total activation to then activate the post-code segment /s/ which is not competing with any other segment. In this way, P3, P4, and P5 can work in concert to produce the output /p/-/l/-/u/-/g/-/z/.

P3: [+obj][+round][+blocks]-----> /p/-/l/-/u/-/g/

P4: [+obj][+soft][+covers][+pl]-----> /r/-/u/-/g/-/z/

P5: [+obj][+anim][+small][+crawls]-----> /b/-/u/-/g/-/z/

Note that both the partial semantic match and top-down feedback from the /u/ and /g/ in P3 contribute to the activation of P4 and P5. The more productions like P4 and P5 that get activated, the stronger the analogy.

3.3 *Combination*

In ACT*, combination occurs when the input causes two or more rules to fire. If we look at the activation of a form like "dogs," we see that it can be produced by either rote or combination. In fact, for regular forms, the two modes of processing converge on the same solution. The rules below would implement the combination route to pluralization:

P6: [+obj][+canine][+furry]-----> /d/-/aw/-/g/

P7: [+pl]-----> /z/,/s/,/i/-/z/

P7 is actually an abbreviation for the following series of productions:

P7′: [+pl]-----> /z/

P7′′: [+pl]-----> /s/

P7′′′: [+pl]-----> /iz/

Of course, since P7′, P7′′, and P7′′′ differ in strength, the elements on the right-hand side of the abbreviated rule P7 also differ in strength.

3.4 *Selections*

Within this same interactive production framework, we can also illustrate the operation of rules governing the selection between alternative segments of morphophonemes. Following Hooper (1976) and Hudson (1980), allomorphic variation is viewed as based on morphophonemic alteration. In the current model, selections on a given level are bound to items on that level. Thus, morphophonological selections are bound to morphemes, phonological selections are bound to segments, and syntactic selections are bound to relational features. Morphophonological selections serve to boost the activation of one of the alternatives in a morphophoneme. Consider the rule for the /f/-/v/ alternation in English plurals. here P10 converts the /f/ to /v/;

P8: [+human][+female][+married]-----> /w/-/al/- coda/f/, coda/v/

P9: [+pl]-----> /-s/ or /-z/ or /-iz/

P10: coda/f/, coda/v/ + [+pl]-----> /coda/v/

In P8, coda/f/ receives stronger activation than coda/v/, so that it would normally be the winner. However, P10 gives /v/ additional activation to overcome the activation of the /f/. Rules like P10 can be stated with a variety of levels of generality by conditioning them on (a) segments, (b) competitions between segments, and (c) morphemes. In languages like Hungarian,

rules such as P10 can be bound to a whole series of alternative affixes and alternative stems, thereby demonstrating a certain limited productivity. Even in English, P10 applies not only to *wife*, but also to *knife, life, leaf*, and so on.

Phonological selections operate much like morphophonological selections, except that they are conditioned on segments rather than morphemes. For example, the change of an /n/ to a /m/ in Hungarian *honve:d---> homve:d* can be controlled by P13 in conjunction with P11 and P12.

P11: coda/n/---> coda[+dental] coda[+voiced] coda[+continuant] coda[+nasal]

P12: onset/v/---> onset[+labiodental] onset[+voiced] onset[+cont.] onset[+fric.]

P13: coda[+dental]+onset[labiodental----> coda[+labial]+onset[+labiodental]

3.5 *Cooperation and Competition*

It is worth noting that the four modes of activation—rote, combination, analogy, and selection—can cooperate or compete. For example, rote, analogy, and combination can each cooperate to activate the plural form /p/-/l/-/u/-/g/-/z/. In cases like this, where all of the rules agree on the items being activated, we do not expect to see any errors. In fact, rules may disagree and we will still see no error, if the strongest rules are also the correct rules. Far more interesting are the cases where disagreement leads to error, since these tell us something about competition processes that are usually masked by correct performance. Three types of errors can be traced directly to competition between lexical items. These include *head-on blends*, a sub-type of blending we call *insertion*, and *semantic extension*.

3.5.1 *Head-on Blends.*

When two competing items fail to suppress each other, they end up activating their respective segments. The problem of resolving the competition then falls on the shoulders of phonological activation. In such cases of fight-to-the-finish, head-on, competition, some segments may come from one morpheme and some from another. Consider this example:

P14: [+gustatory][+property]. . .-----> f l a v o r

P15: [+gustatory][+property]. . .-----> t a s t e

If these two rules are approximately equally active, we can get a mixture of their two actions such as *flaste*. In the current model, each of the segments in these productions is actually being given a full autosegmental positional characterization. Thus, both the /fl/ of *flavor* and the /t/ of *taste* are defined as onsets of the first syllable. Thus, these two segments are competing with each other. Similarly, the /st/ of *taste* is competing with the /v/ of *flavor*. The /r/ of *flavor* is competing with an empty second syllable in *taste*. For brevity, these syllable/position/slot characterizations of segments are being omitted.

Blends occur when two items are competing for the same slot and neither can achieve dominance. Children make blends just as adults do (MacWhinney, 1974; Leopold, 1937–1949), indicating that head-on competition between content words is fairly basic to the system. Examples of blends from Stemberger (1982) include: Stiff—stuck—stuff it in this box. (stick/stuff); it removes the dirst. (dirt/dust); Look at him proon himself. (groom/preen).

3.5.2 Insertion. Competition can also lead to the insertion of a whole morpheme into a position where it is not directly competing with other morphemes. In general, this means that closed-class morphemes can intrude themselves in this way, whereas competition between open-class morphemes, which function as the nuclei for lexical arenas, leads to head-on blending. Adult errors of this type, with the insertions marked by parentheses, include *ingredients(es), seven-day-(year)-old baby, landed(ed), you needed (a) bilateral damage, we just put it in (with) the cage,* and *I've been (keep) thinking that*. Children do a lot of this too. English examples include *foots(es), sevens(es), mommy get (it) ladder,* and *me and Ross (and I)*. Some Hungarian examples are *pingvink(ek)* and *ram(om)(ra)*. A French example is *mon (mien de) chapeau (a moi)*. In general, insertions occur when the competing units are able to open up structures that are not competing for the same positional slot. When competing structures are targeted for the same position, we will find blends rather than insertions.

3.5.3 Semantic Extension. Sometimes, the wrong lexical item can be retrieved on the basis of a partial match. For instance, consider the following overlapping rules:

P16: [+covering][+poss-animal][+strands]-----> hair

P17: [+covering][+poss-animal][+barbs]-----> feather

P18: [+covering][+head-loc][+strands]-----> hair

Rules P16 and P18 are two rules for *hair*, one for animals in general and one for the top of human heads while P17 is a rule for feather. One might assume that, if one's goal was to describe feathers on the head, P16 and P18 might in combination overrule P17 and lead to the production of *hair*. If this happens, it would be a *semantic extension*. Semantic extensions include case errors, antonym substitutions; within-class errors, errors based on perceptual similarity, functional errors, and so on. Although semantic extensions are not rare in adults, a truly huge number of semantic extensions have been reported from children. In deliberate attempts to elicit extensions, researchers have found up to 40% extensions. Even in free-speech corpora, very high levels have been reported from some children.

3.6 *Top-down Support*

In comprehension, top-down support is spread from the meanings being activated back down to the morphemes matching those meanings. In generation, activation spreads from the sound segments being activated back up to the morphemes matching the output. Thus, the equivalent of top-down support in generation is "bottom-back" support. Three types of speech errors indicate that items that are activated for production feed activation back up to the morphemic and semantic levels. These three types are: semantic extensions within idioms, malapropisms, and haplologies.

3.6.1 Semantic Extensions Within Idioms. For semantic extensions to occur within idioms, there must be some top-down support. This is because, in the context of the idiom, the word being replaced does not have its usual meaning. Examples of such errors are *I'm in the store for a parakeet* for *I'm in the market for a parakeet* and *It was skinnily disguised as...* for *It was thinly disguised as...* The idiom can be represented by a rule like P19:

P19: [+state][+description][+desire][+purchase]---> "be" "in" "the" "market" "for"

Here the words in quotes are taken to be concepts or prepackagings of semantic material. The item "market" reactivates its component pieces by P20:

P20: "market"---> [+public[+place][+purchasing]

These features then give top-down support for "store" which replaces "market" in the output. Semantic extensions within idioms have never been reported for children.

3.6.2 Malapropisms. Further evidence for top-down support comes from klang-associations or malapropisms such as *excavator* for *executor*. For further examples see the footnote. Examples from Stemberger (1982) include *you used coupons* for *you used croutons, Rape! Murder! Plunge!* for *Rape! Murder! Pillage!, There was a meeting in Toyota* for *There was a meeting in Kyoto*, and *of getting 'em sick to test their reproductive sickness* for *of getting 'em sick to test their reproductive success*. There are very few reports of true malapropisms in children. It is true that children make errors like *Judas Asparagus* for *Judas Iscariot*, but in children such errors are usually due to misperception of some new word.

3.6.3 Haplology. Consider the following set of rules:

P21: [+desire]...-----> /n/-/i/-/d/

P22: [+past]-----> /t/, /d/, or /i/-/d/

Acting directly, P21 and P22 would lead to the correct form *needed*. However, in cases where P22 failed to fire in time or with sufficient strength, the principle of top-down support would allow the /d/ in /n/-/i/-/d/ to be matched to the /d/ in P22. This would then "use up" the past feature, according to the principle of data refractoriness, and block additional application of P22. The output in this case would be *need*, rather than *needed*. Both adults and children produce many haplologies of this sort and languages have developed a variety of specific ways of dealing with this pervasive tendency to haplology (Menn & MacWhinney, 1984).

4. SYNTACTIC PROCESSING

What was described in the previous section was a model of how lexical items are generated and how lexical errors arise in children and adults. This account based on the work of MacWhinney. A recent detailed update of that work is given in MacWhinney (in press). What we will describe in this section is Anderson's theory of syntactic control of sentence production. We can think of syntax as a hierarchical control system that sequentially orders the lexicalization of individual lexical items. The architectural principles are identical to those we used for the lexical component. Anderson's theory is given a fairly thorough exposition in Chapter 7 of Anderson (1983). Here we will just present an overview of its main points.

4.1 Free and Bound Rules

As with the lexical component, there are a set of rules which decompose higher level goals into lower-level goals. Although the principles governing rule structure are similar, syntactic rules make far more extensive use of hierarchical control than do lexical or morphological rules. Consider this rule:

P23: IF the goal is to describe a proposition involving a relation between an agent and an object
THEN set as subgoals to:
1. lexicalize the agent,
2. lexicalize the relation,
3. lexicalize the object.

This rule underlies the canonical subject-verb-object order of active English sentences. Such a rule sets subgoals corresponding to the phrase structure of language. Basically, phrase structure is the goal structure produced by syntactic productions. General rules like P23 make use of conditions based on semantic features such as [+agent] and [+object]. Rules of this type are what MacWhinney (1982) has called "lexically-free" syntactic rules.

Many rules make mention of particular lexical items. MacWhinney (1982) treats these as "lexically bound." Examples of such rules include:

P24: IF the goal is to lexicalize a constituent and the listener has heard the constituent mentioned,
THEN set as subgoals to:
1. activate "the"
2. lexicalize the constituent

P25: IF the goal is to lexicalize an item and the item is a unitary concept
THEN set as subgoals to:
1. activate the sound of the item

P26: IF the goal is to describe a relation and it is an ongoing action and the action is currently happening
THEN set as subgoals to:
1. activate "is"
2. lexicalize the action
3. activate "ing"

Rules like P25 above provide the bridge between the syntactic component and the lexical component by achieving lexical look-up. This provides the index to lexical productions like P1 and P2. Together these rules can generate a sentence like *The dog is chasing the cat*. Note that rules like P26 provide for control of discontinuous elements by goal embedding. Interfacing such productions with a parallel mechanism requires the marking of places in syntactic rules where "pushes" are made to separate lexical arenas. When processing is complete in a given lexical arena, a "pop" occurs, and the material in that arena is then sent back to the matrix production for compilation.

4.2 Transformations

Transformations can be understood as operations on goal trees such as the one formed by the firing of P23-P26. First, planning productions operate to produce a goal structure, and then transformations reorganize them. The following is an example of the planning production that would underlie the generation of the question *Is the dog chasing cat?*

P27: IF the goal is to question the truth of a proposition
THEN set as subgoals to:
1. plan the communication of the proposition
2. move the element coding tense to the beginning
3. execute the plan

This production would plan the generation of the declarative sentence, then transform it to question form, and then actually produce the transformed sentence. Thus, ACT*s planning and transformation processes underlie the transformational component of English.

We realize that this formulation of transformations is much different from that in recent versions of standard theory, where a single transformation is allowed to move anything and where filters and constraints limit the varieties of derived structures. We agree with the emphasis on stating constraints on the grammar. (In fact, the rule types we allow are limited to lexical mapping rules like P1, selection or allomorphy rules like P10, and three types of syntactic rules: lexically-bound rules, free rules, and class rules.) However, transformations occupy a somewhat different role within our psycholinguistic approach, since we see transformations are operating in a "metaprocedural" fashion upon the basic output of the grammar (as in subgoal 1 of P 27).

4.3 *Compilations of Transformations*

For even fairly complex structures, speakers are able to generate output directly from phrase structure without utilizing explicit on-line transformations (Bresnan, 1982). In both the dialectic model and ACT* this occurs as the result of a compilation process which can "compile out" the explicit planning and transformation. Compilation of one version of the question transformation would produce a rule of the form:

P28: IF the goal is to question the truth of a proposition between an agent and an object and the relation describes an ongoing action and the action is currently happening
THEN set as subgoals
1. To activate "is"
2. To lexicalize the agent
3. To activate the name for the relation
4. To activate "ing"
5. To lexicalize the object.

In this case, information controlling the activation of the verb auxiliary "is" has migrated to the main generation production. The previous production had planned a sentence, transformed it, and then executed it; the current production directly generates the question. Such a compiled production makes possible more automatic and less resource-demanding sentence generation. Forming such compiled productions is one of the nine acquisitional modes we will discuss later.

4.4 *Syntactic Errors*

Anderson (1983) ignored the impact of interactive activation in his exposition of the structure of the syntactic component. Without considering interactive effects, it is difficult to account for certain types of speech errors in which whole morphemes are either lost or misplaced. However, the current formulations of production action and competition can be used to explain both morpheme loss and morpheme misplacement. In both children and adults, closed-class items are lost far more frequently than open-class items, despite the fact that closed-class items are more frequent and hence inherently stronger. We view the loss of closed-class items as a result of the binding problem that arises during the transfer from serial to parallel control.

Morpheme misplacements include transpositions, anticipations, and perseverations. The mechanism underlying each of these types is the same. If there are two simultaneous goals to activate two lexical terms, it is possible to have the description of one be invoked through partial matching in the context for the other. This will produce transpositions with affix stranding, as in these examples from Stemberger (1982): *he doesn't have any closets in his skeleton* for *he doesn't have any skeletons in his closet*, or without stranding, as in *there's a beard hanging from your hair* or *there's a hair hanging from your beard*. Anticipations such as *I got a paper on my test* for *I got an A on my paper* or *how good is that for?* for *how long is that good for?* seem to begin as transpositions. However, when the speaker reaches the point where the second word is being produced, s/he inserts an associate of what was originally the second element, rather than inserting the original first element. Perseverations, on the other hand, are more likely to involve exact repetitions as in *the bed was lying on the bed* for *the bed was lying on the book* or *raccoons sometimes eat raccoons* for (raccoons sometimes eat frogs).

Morpheme loss occurs most often when a closed-class item does not reach selection threshold by the time its open-class item is read. Examples of morpheme loss with the missing items enclosed in parentheses include: *I just wanted to (ask) that*, *the points distribute the (electricity) out*, *you don't have to worry (about) that*, *see what (the) child does*, *one strategy for do(ing) everything*, and *he relax(es) when you go away*. Children also leave out items. However, it is difficult in child data to distinguish loss from non-acquisition.

5. *MONITORING*

So far, we have described how the lexical and syntax components work. Now we turn to questions of learning and development. Specifically, how are these syntactic and lexical rules acquired? A crucial claim in the dialectic model is that learning proceeds on the basis of identification of some error, i.e., some mismatch between the child's system and new data s/he is processing. Monitoring looks for two things: completeness and accuracy.

5.1 *Completeness Monitoring*

When people are talking, they can check to see if what they have said fails to include a part of what they wanted to say. When we detect such omissions, we may retrace or restart our utterances. However, if we have not yet acquired the needed item, we must simply confirm our desire to learn it. When people are listening, they can check to see if there is material in the verbal message that is unfamiliar. In either case, the language user is able to detect the existence of new and unfamiliar lexical items. By accepting partial regularities (MacWhinney, 1983; Menn and MacWhinney, in press), speakers indicate that they do, in fact, check to see if what they are going to say includes all of what they intended to say.

5.2 *Accuracy Monitoring*

Even when monitoring detects no mismatch between the verbal message and the situation, it may be able to detect formal inadequacies in the verbal message. This can occur in either production or comprehension. In production, the generation of a combinatorial or analogic form can facilitate retrieval of a weak rote item. Children can then check to see if there is a discrepancy between the weak rote receptive form and the combinatorial expressive forms. This mode of monitoring turns out to be quite effective for morphophonological acquisition (MacWhinney, 1978). However, applying accuracy monitoring above the level of the word requires that there be long rote units. In comprehension, these length limitations may be relaxed. While listening to incoming sentences, the language learner uses his system to generate his own versions of the target sentences. If these match the actual input, no error is present. If they do not match, and if the child places trust in the input, s/he must then pass the mismatch on to acquisition. This mode of monitoring is extremely powerful, since it provides acquisition with both new formal structures and interpretations of those structures. When accuracy monitoring works in comprehension, it can allow the child to outgrow even persistent, self-taught errors. Monitoring during comprehension is the main mode of monitoring in the ACT* stimulation of language acquisition.

6. *ACQUISITION*

In acquiring language, the child's task is to formulate a set of rules that can produce sentences appropriate to the speech community. Some of these rules are lexical rules and some are syntactic rules. Some are highly general and some are quite specific. In this section, we will examine a set of nine acquisitional strategies which, when taken together, can allow the language learner to acquire the lexical and syntactic rules of his language. These strategies are linked together in terms of the overall dialectic.

6.1 *Lexical Rules: Amalgam Acquisition*

When monitoring for completeness, the child may note a spatiotemporal contiguity between a phonological cluster produced by another speaker and some meaning cluster. When this occurs, the child can associate the phonological cluster to the meaning cluster as a new lexical item. What is important here are the principles that govern clustering or segmentation on the phonological and the semantic levels. On the phonological side, the claims are (a) that clustering is determined by intensity, pitch, and juncture; (b) that recent segments and stressed segments are stored most clearly; and (c) that presentation of a single intonational group in isolation enhances acquisition by precluding any need for segmentation. The idea of early words as phonological clusters is supported by a variety of acquisitional phenomena discussed in MacWhinney (1978):

1. When morphemes appear in several allomorphic shapes, the child tends to take the citation form as basic. This suggests that children are picking up words from utterances where the word is being produced by itself.
2. Affixes are acquired quite early, earlier than adpositions, but their first uses indicate that they function as integral parts of the stems with which they were learned. (See the discussion of rote forms in section 3.1.)
3. The importance of phonological packaging is reflected in the fact that children tend to preserve stressed syllables and final syllables in newly learned words.

On the semantic side, the assumption is that children are attempting to acquire words for meanings that they want to express and for which they think there ought to be a word. Clark (1982) calls this the principle of "conventionality" and holds that the child realizes that "for certain meanings, there is a conventional word or word-formation device that should be used in the language community."

6.2 *Inflectional Rules: Component Analysis*

When monitoring for completeness, the child may note that a word contains noncomprehended segments. If situational clues are present, the child can link this noncomprehended section to the semantic clue. This analysis of components out of amalgams makes the default assumption that words are compositions of continuous pieces. A variety of reports indicate that, in fact, reanalysing his own data from Mayan, as well as Brown's (1973) data on English, Pye (1979) has shown that the order of acquisition of grammatical morphemes in both English and Mayan is better predicted by ease of segmentability than by either semantic complexity or frequency.

6.3 *Syntactic Rules*

When monitoring for completeness, the child may understand, say, two words, but not have a rule to govern their syntactic relation. When this occurs, the child forms a new rule on both the free and bound levels. Rules on the free level must be formulated in terms of already available features and units. Rules on the bound level are formulated in terms of lexical items. Finally, in ACT*, rules may be formulated in terms of formal syntactic classes. For a complete discussion of these three levels of syntactic rules, details on the structure of each, and the evidence supporting the importance of each type, please consult MacWhinney (1982).

6.4 *Strengthening*

Strengthening serves to identify the correct rules and increase their probability of application. Every time a rule or item is used during either expression or reception and no error results, it gains in strength. Whenever a rule or item applies unsuccessfully, it loses strength. Frequency of correct application is an important factor in accounting for the development of both items and rules. In regard to lexical items, MacWhinney (1978) has shown that frequent forms tend to resist morphophonological overregularizations. Thus, in both children and adults, errors like *goed* are relatively less frequent than errors like *felled*. In languages that have extensive allomorphy, it is generally true that the allomorph that is overgeneralized is the one which is most frequent and, hence, strongest. There are now a total of 14 examples from different languages supporting this observation. In regard to rules, MacWhinney (1978) demonstrated that the order of acquisition of 15 rules in Hungarian morphophonology was closely correlated with their relative order of frequency of correct application. In other languages, detailed analyses of this type have not yet been conducted, but it has been reported that rules affecting affixes are generally acquired before rules affecting stems, apparently because of the fact that individual affixes are much more frequent that individual stems. In the area of semantics, similar effects have been reported. For example, the most frequently overextended polyseme of the Hungarian denominative verbalizer *-ol* is the agentive polyseme, which is also the most frequent in the language. In syntax, the complement structure for common verbs like *tell* is overgeneralized to replace that of less frequent verbs like *promise* (Chomsky, 1969).

6.5 *Generalization*

Generalization collapses many specific rules into a single general rule. For example, the child may have the adjective "good" encoded in two forms, one for the singular as in "good girl" and one for the plural as in "good girls". When the child realizes that these two forms are equivalent, s/he achieves a generalization. If this is done generally across all adjectives, many forms will be merged into a single rule. To take another example, in Hungarian the rule for final-a-lengthening may initially be stated to apply only when the suffix is the plural /-Vk/. When the child realizes that the rule applies with all suffixes that begin with vowels or deletable vowels, s/he achieves a significant generalization.

6.6 *Discrimination*

Discrimination adds extra features to overly general rules. For instance, if there is a rule that activates the same verb form in present or past, two different versions of it might be produced, one with a test for present tense and the other with a test for past tense. Whereas in generalization the focus is on the detection of extraneous conditions in rules, in discrimination the focus is on missing conditions in rules.

6.7 *Proceduralization*

Anderson, Greeno, Kline, and Neves (1981) underscore the importance of (pre)compilation of highly specific rules in achieving accuracy and automaticity. One compilation process is called proceduralization. It takes a general rule and makes it specific to a particular circumstance of application. So, for instance, one can imagine the following lexical rule:

P29: IF the goal is to communicate a feature set and a sound-structure has been used to communicate that feature set
THEN set as subgoals to communicate the components of that sound-structure.

Thus, if the speaker recalls hearing /d-a-d/ to communicate {[+parent] [+male]} this would lead to the subgoals of generating /d/, /a/, and /d/. However, through practice proceduralization will form a version of the rule specific to this lexical item:

THEN set as subgoals to:
1. activate /d/ in onset position
2. activate /a/ in nucleus position
3. activate /d/ in coda position

Because all of the subgoals here involve lexical activation, proceduralization can make P30 fully parallel.

6.8 *Composition*

The other compilation process is composition. It combines the effect of a number of rules into a single rule. Again, it occurs with practice. Thus, P23 and P26 given earlier might be composed into the following rule:

P23: IF the goal is to describe a proposition involving a relation between an agent and an object and the relation describes an ongoing action and the action is currently happening
THEN set as subgoals to:
1. lexicalize the agent,
2. activate "is"
3. lexicalize the action
4. activate "ing"
5. lexicalize the object

The effect of both proceduralization and composition is to make rapid, special-purpose rules. They serve to make the process of language generation more and more automatic. They also help the system in recovering from overgeneralization. Butler and MacWhinney (in press) have noted that composition may lead the child into self-taught errors. For example, composing attempts to add the regular past suffix to irregular verbs will lead the child to learning forms like **bringed* and **catched* by rote.

6.9 *Inference*

Bates and MacWhinney (1982) stressed the importance of functional characterizations of syntactic classes. There is evidence that children not only are able to infer the class of a word from cooccurrence data, but even can override semantics in placing items into classes. For example, Katz, Baker, and Macnamara (1974) found that, beginning around 17 months, girls who were given a proper name for a doll learned this name better than girls who were given a common noun. In the proper noun frame, girls were told that the doll was called "Zav"; in the common noun frame they were told that the doll was "a zav." Thus, even at this early age, children seem to realize that names with articles are common nouns and names without articles are proper nouns. This ability to infer the semantics of words on the basis of cooccurrence continues to develop. By age 8, Werner and Kaplan (1950) were able to show in their classic "corplum" experiment that children could acquire many aspects of the semantics of abstract nouns from highly abstract sentence contexts.

The degree to which conjugation, declension, part-of-speech, and grammatical role should be understood as formal rather than functional categories continues to be a topic of much investigation (Bates & MacWhinney, 1982). Most parties (e.g., Maratsos, 1982; Pinker, 1982) agree that categories are defined initially in functional terms. Even highly arbitrary systems such as German gender have been shown to have a partially functional basis (Kopcke & Zubin, in press). The ACT* model in Anderson (1983) describes a generalization scheme for forming such categories that can use both semantic and formal features.

7. *CONCLUSION*

The current model incorporates a number of important advances over our earlier independent efforts. Most improved is the theory of lexical activation, which now provides a unified account of the interaction of rote, analogy, combination, and selection while also dealing with the various types of speech errors. We next hope to focus our attention on a theoretical unification of the nine acquisitional strategies that we have just outlined. We believe that a similar type of theoretical unification can be achieved here too.

In this paper, our focus has been on the construction of an account of language development that is maximally grounded in the theories of cognitive psychology and linguistics. By emphasizing the connections between language structures, language processing, general cognitive processing, and general learning mechanisms, we believe that it is possible to construct a coherent account of the great mass of data that has been collected in the last 20 years of research in child language acquisition. We believe that unitized approaches of this sort avoid the "cognitive ugliness" of which Roger Brown has warned us. We also believe that the unified approach we have sketched out can also form an important part of the general theories of both psychology and linguistics.

REFERENCES

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. (1977). Induction of augmented transition networks. *Cognitive Science, 1,* 125-157.

Anderson, J. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85,* 249-277.

Anderson, J. (Ed.). (1981). *Cognitive skills and their acquisition.* Hillsdale, NJ: Erlbaum.

Anderson, J. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J., Greeno, J., Kline, P., & Neves, O. (1981). Acquisition of problem-solving skill. In J. Anderson (Ed.), *Cognitive skills and their acquisition.* Hillsdale, NJ: Erlbaum.

Baker, C. (1977). Comments on the paper by Culicover and Wexler. In P. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal syntax.* New York: Academic Press.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art.* New York: Cambridge University Press.

Bierman, A. W., & Feldman, J. A. (1972). A survey of results in grammatical inference. In S. Watanabe (Ed.), *Frontiers of pattern recognition.* New York: Academic Press.

Block, H., Moulton, J., & Robinson, G. (1975). Natural language acquisition by a robot. *International Journal of Man-Machine Studies, 7,* 571-608.

Blum, L., & Blum, M. (1975). Toward a mathematical theory of inductive inference. *Information and Control, 28,* 128-155.

Braine, M. D. S. (1971). The acquisition of language in infant and child. In C. Reed (Ed.), *The learning of language.* New York: Appleton-Century-Crofts.

Bresnan, J. (Ed.). (1982). *The mental representation of grammatical relations.* Cambridge, MA: MIT Press.

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

Brown, R. (1977). *Word from the language acquisition front.* Invited address at the Eastern Psychological Association, Boston.

Butler Platt, C., & MacWhinney, B. (1983). Error assimilation as a mechanism in language learning. *Journal of Child Language, 10,* 401-414.

Chandrasekaran, B., & Reeker, L. H. (1974). Artificial intelligence—a case for agnosticism. *IEEE Transactions on Systems, Man and Cybernetics, 4,* 8-94.

Chomsky, C. (1969). *The acquisition of syntax in children from 5 to 10.* Cambridge, MA: M. I. T. Press.

Chomsky, N. (1980). On binding. *Linguistic Inquiry, 11*, 1–46.
Clark, E. (1982). Language change during language acquisition. In M. Lamb & A. Brown (Eds.), *Advances in Developmental Psychology* (Vol. 2). Hillsdale, NJ: Erlbaum.
Culicover, P., & Wexler, K. (1977). Some syntactic implications of a theory of language learnability. In P. Culicover, T. Wascow, & A. Akmajian (Eds.), *Formal syntax*. New York: Academic Press.
Dresher, B. E., & Hornstein, N. (1976). On some supposed contributions of artificial intelligence to the scientific study of language. *Cognition, 4*, 321–398.
Feldman, J. (1972). Some decidability results on grammatical inference and complexity. *Information and Control, 3*, 244–262.
Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk*. London: Academic Press.
Gold, E. (1967). Language identification in the limit. *Information and Control, 10*, 447–474.
Hamburger, H., & Wexler, K. N. (1973). Identifiability of a class of transformational grammars. In K. J. J. Hintikka, J. Moravcsik, and P. Suppes (Eds.), *Approaches to natural language*. Dordrecht-Holland: D. Reidel.
Hamburger, H., & Wexler, K. (1975). A mathematical theory of learning transformational grammar. *Journal of Mathematical Psychology, 12*, 137–177.
Harris, L. (1977). A system for primitive natural language acquisition. *International Journal of Man-Machine Studies, 9*, 153–206.
Hinton, G., & Anderson, J. (1981). *Parallel models of associative memory*. Hillsdale, NJ: Erlbaum.
Hooper, J. (1976). *An introduction to natural generative phonology*. New York: Academic Press.
Horning, J. J. (1969). *A study of grammatical inference* (Tech. Rep. No. CS 139). Palo Alto, CA: Stanford University, Computer Science Department.
Hudson, G. (1980). Automatic alternations in non-transformational phonology. *Language, 56*, 94–125.
Katz, N., Baker, E., & MacNamara, J. (1974). What's in a name? A study of how children learn common and proper names. *Child Development. 45*, 469–473.
Kelley, K. L. (1967). *Early syntactic acquisition*. (Tech. Rep. No. P-3719). Santa Monica, CA: Rand Corporation.
Kiss, G. (1973). Grammatical word classes: A learning process and its simulation. In G. Bower (Ed.), *The psychology of learning and motivation*. New York: Academic Press.
Klein, S., & Kuppin, M. (1970). An interactive heuristic program for learning transformational grammars. *Computer Studies in the Humanities and Verbal Behavior, 3*, 144–162.
Kopcke, K. & Zubin, D. (in press). Zur Frage der psychologischen Realitat von Genuszuweisenden Regeln zu den einsilbigen Nomen der deutschen Gegenwartssprache. *Linguistische Berichte*.
Langley, P. (1982). Language acquisition through error recovery. *Cognition and Brain Theory, 5*, 211–155.
Leopold, W. F. (1937–1949). *Speech development of a bilingual child: A linguist's record* (Vols. 1–4). Evanston, IL: Northwestern University Press.
Levelt, W. (1975). *Formal grammars in linguistics and psycholinguistics*. The Hague: Mouton.
MacWhinney, B. (1974). *How Hungarian children learn to speak*. Unpublished doctoral dissertation, University of California, Berkeley.
MacWhinney, B. (1975). Pragmatic patterns in child syntax. *Stanford Papers And Reports on Child Language Development, 10*, 153–165.
MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development, 43*, Whole No. 1.
MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (Ed.), *Language acquisition: Vol. 1. Syntax and semantics*. Hillsdale, NJ: Erlbaum.

MacWhinney, B. (1983). Miniature linguistic systems as tests of universal operating principles. *Journal of Psycholinguistic Research, 12*, 467–478.
Maratsos, M. (1982). The child's construction of grammatical categories. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art*. New York: Cambridge University Press.
McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of the basic findings. *Psychological Review, 88*, 375–402.
McMaster, I., Sampson, J., & King, J. (1976). Computer acquisition of natural language: A review and prospectus. *International Journal of Man-Machine Studies, 8*, 367–396.
Menn, L., & MacWhinney, B. (1984). The repeated morph constraint: Topward an explanation. *Language. 60*, 519–581.
Miller, G. (1967). *The psychology of communication*. New York: Basic Books.
Newell, A., & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
Pao, T. W. L. (1969). *A solution of the syntactical induction-inference problem for a non-trivial subset of context-free languages* (Tech. Rep. No. 70-19). University of Pennsylvania: The Moore School of Electrical Engineering.
Pinker, S. (1979). Formal models of language learning. *Cognition, 7*, 217–283.
Pinker, S. (1982). A theory of the acquisition of lexical-interpretive grammars. In J. Bresnan (Ed.), *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
Pye, C. L. (1979). *The acquisition of grammatical morphemes in Quiche Mayan*. Unpublished doctoral dissertation, University of Pittsburgh.
Reeker, L. (1976). The computational study of language acquisition. In M. Yovits & M. Rubinoff (Eds.), *Advances in computers*. New York: Academic Press.
Siklossy, L. (1972). Natural language learning by computer. In H. A. Simon & L. Siklossy (Eds.), *Representation and meaning: Experiments with information processing systems*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
Stemberger, J. (1982). *The lexicon in a model of language production*. Unpublished doctoral dissertation, University of California, San Diego.
Werner, H., & Kaplan, E. (1950). Development of word meaning through verbal context: An experimental study. *Journal of Psychology, 29*, 251–257.
Wexler, K. & Culicover, P. (1980). *Formal principles of language acquisition*. Cambridge, MA: MIT Press.
Wexler, K., Culicover, P., & Hamburger, H. (1975). Learning-theoretic foundations of linguistic universals. *Theoretical Linguistics, 2*, 215–253.
Wexler, K. N., & Hamburger, H. (1973). On the insufficiency of surface data for the learning of transformational languages. In K. J. J. Hintikka, J. Moravesik, and P. Suppes (Eds.), *Approaches to natural language*. Dordrecht-Holland: D. Reidel.
Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.
Woods, W. (1970). Transition network grammars for natural language analysis. *Communications of the ACM, 13*, 591–606.