The (il)Logical Problem of Language Acquisition

Brian MacWhinney
Department of Psychology
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
brian@andrew.cmu.edu

Abstract

The fact that children often appear to learn little from corrective feedback has led theorists to construct the "logical problem of language acquisition" (LPLA). The idea is that, without further formal constraints, language can be formally proven to be unlearnable. This paper argues that the LPLA is based on a restricted view of the nature of language, the nature of the learner, and the nature of the learning environment. When we examine the full set of forces channeling language learning, including competition, recasts, conservatism, and indirect negative evidence, we see that language is in fact highly learnable and the LPLA is not well motivated. In its place, we hope that language theory can focus on the analysis of overgeneralizations as ways of diagnosing the shape of underlying analogic pressures.

Introduction

No other topic has received more attention in recent theories of language and language learning than the role of corrective feedback or **negative evidence** (Pinker, 1984; Wexler & Culicover, 1980). The so-called "logical problem of language acquisition" (LPLA) (Baker, 1979) has become a major conceptual pillar underlying a vast amount of current work in generative linguistics (Gass & Schachter, 1989; Matthews & Demopoulos, 1989; Weissenborn, Goodluck, & Roeper, 1992) and has played a major role in the development of the "principles and parameters" view of linguistic competence (Chomsky, 1981). Within psycholinguistics, "Learnability Theory" (Pinker, 1984) has used the LPLA as the fundamental way of motivating a nativist approach to linguistic theory. Various proposed solutions to the LPLA currently occupy the focus of much of linguistic theory, language development, crosslinguistic analyses, and cognitive development. These solutions have potential significance for cognitive science as a whole. Despite the centrality of this topic, there has been relatively little public discussion of the assumptions underlying the standard analysis of the LPLA. This paper analyses the standard assumptions and calls into question the extent to which the presumed non-availability of negative evidence constitutes a logical problem.

Gold's Problem

In a 1967 paper entitled "Language Identification in the Limit" , Gold (1967) examined the learnability of grammars on different levels of the Chomsky hierarchy (Chomsky, 1963) He contrasted two different language learning situations: **text presentation** and **informant presentation**. With informant presentation, the language learner can receive feedback from an infallible informant regarding the grammaticality of each and every sentence generated during the process of language learning. Whenever the learner formulates an overgeneral guess about some particular linguistic structure, the informant labels the resulting structure as ungrammatical and the learner uses this information to restrict the developing grammar. In the case of text presentation, the learner only receives information on acceptable sentences and no information regarding ungrammaticality is available.

With text presentation, if the learner formulates an overly general hypothesis, there is no way to recover from the overgeneralization and the grammar remains forever off the mark. Consider a very simple example. The learner receiving text presentation can use a corpus of regular verbs to induce the "add -ed" rule which will then produce the overgeneralized form "goed". Without information regarding the ungrammaticality of "goed", the learner will never be able to recover from this overgeneralization.

Gold showed that this would be true for the learning of all but the simplest forms of language. If the set of languages being explored includes only finite languages (languages generated by regular, Markovian processes), text

presentation is adequate. But if the set of possible languages includes at least one non-finite language, the learner may rush into guessing that language. Once he does, there will be no way to retreat from the incorrect guess, even if the non-finite language is too general and produces some strings that are ungrammatical.

Horning (1969) showed that Gold's problem can be avoided if one relaxes the criterion for identification in the limit, allowing for a probabilistic approximation to the grammar. In this case, a wider variety of context-free languages as well as some context-sensitive languages can be learned on the basis of only text presentation.

If linguists were willing to believe that language could be described by finite-state automata, or if they were willing to accept probabilistic identification as a correct description of language learning, then Gold's analysis would not present a major empirical challenge. However, it seems unlikely that the full scope of human language can be described by a finite-state automata. Given this, and given the attempt to view language identification as deterministic, many psycholinguists have accepted Gold's analysis and assumed that it provides a solid basis for the LPLA.

Overt Correction

Gold's analysis leaves open the question of the empirical status of text presentation and informant presentation. The data required for text presentation are clearly available, since language learning involves massive exposure to acceptable sentences. But what about the unacceptable sentences required for "informant presentation"? Intuitively, it makes sense to imagine that the parent is capable of telling the child that particular sentences or strings are ungrammatical. But how would this be done in practice?

One system for informant presentation would involve full overt correction. For example, if the child says "I go to the store", the parent could conceivably say, "It is not grammatical to say 'I go to the store'. You must say 'I am going to the store'." Studies of interactions between children and their parents (Bohannon, MacWhinney, & Snow, 1990; Bohannon & Stanowicz, 1988; Brown & Hanlon, 1970; Demetras, Post, & Snow, 1986; Farrar, 1992; Hirsh-Pasek, Trieman, & Schneiderman, 1984; Morgan & Travis, 1989; Penner, 1987; Sokolov, 1990) show that sometimes parents do just this. But they only provide full overt correction on rare occasions. Moreover, children often appear to ignore these corrections, persisting for awhile in their old erroneous ways. To give just one example of this type of interaction, consider this dialogue reported by McNeill (1966):

Child: Nobody don't like me.
Mother: No, say "Nobody likes me."
Child: Nobody don't like me.
[dialogue repeated eight times]
Mother: Now listen carefully, say "Nobody likes me."
Child: Oh! Nobody don't likeS me.

Such examples show that overt correction is often initially unsuccessful. However, there is sometimes evidence of the effects of the correction several minutes later.

Complex Feedback

In the Gold framework, what is important is not the shape of the correct form, but rather the information that a form or string is incorrect. Transcript-based research has shown that ungrammatical sentences elicit a pattern of responses that is different from that for grammatical sentences. These differences involve several dimensions of feedback:

1. contingent queries such as "what?" or "who?",
2. requests for repetition such as "excuse me" or "huh?",
3. recasting of the sentence in a more grammatical form,
4. topic shifting, or
5. proxemic, gestural cues signaling miscomprehension.

Together, these responses indicate that the parent's main goal in providing feedback to the child is not the provision of negative evidence, but the extraction of the child's meaning and the maintenance of a successful interaction. When one thinks a bit about the language learning process, this makes sense. If the parent started off from the beginning by providing uniform negative feedback to all ungrammatical sentences, virtually all of the child's first 1000 utterances would be marked with an asterisk and an eyebrow raise. The child would learn little from this process except perhaps to avoid communicating with a person who provides nothing but raised eyebrows. Even if the parent recognizes the importance of identifying ungrammatical sentences, they must check their corrective impulses until the child is able to focus on specific constructions that require detailed feedback.

Given the complexity of the situation facing the child and the parent, it makes sense to

imagine that negative evidence is provided differentially across contexts and constructions. If one looks across the board at all sentences and constructions, one finds that parents are more likely to recast ungrammatical sentences, particularly if those sentences contain only one or two errors (Bohannon & Stanowicz, 1988). However, if the child were to treat this feedback as a simple on-off signal of grammaticality, there would have to be immense amounts of data for each individual construction in order to distinguish signal from noise (Marcus, 1993). It is unlikely that negative feedback is used in this simple on-off way.

Blocking

The major thesis of this paper is that feedback is used by parents not to provide the asterisks required in the Gold framework, but to provide positive evidence to compete with overgeneralizaed forms. To explain this position, let us begin by considering the simplest of cases: the direct blocking of an morphological overregularization by the corresponding irregular form.

Children often produce simple morphological overgeneralizations such as "goed" for "went" or "mouses" for "mice". If learning is not based on corrective feedback, how can it be that they slowly stop producing these overregularized forms (MacWhinney, 1978; Marcus, et al., 1992)? A universally-accepted solution to this particular problem focuses on the idea that "went" **blocks** the production of "goed". Whenever an incorrect form directly competes with a correct form and a complete meaning equivalence can be established between the two forms, we can say that there is a **benign** overgeneralization (Baker, 1979; Baker & McCarthy, 1981) that can be eliminated through blocking. The standard view holds that benign overgeneralizations can be eliminated without any reliance on negative evidence. By contrast, **embarassing** overgeneralizations are those which require negative evidence.

Baker (1979) assumed that production of "goed" could be blocked by a principle of rule-ordering that places specific rules before general rules in the rule cycle. Even if one were to accept the notion that all rule systems are ordered in this extremely strict fashion, Baker's account would still leave open two major questions. The first is how it can be that overgeneralizations of forms like "goed" could arise in the first place. Before the production of the first overgeneralizations, the child is already correctly producing forms like "went" and "fell". If there were no need to learn to place the rule for "went" before the general rule for "goed", these errors would never arise. But once they do arise, there would have to be a learning mechanism that could guarantee that the child would learn to place these two rules in the correct order. However, it appears that learning the ordering of these rules reduces directly to the problem identified by Gold in the first place and Baker's analysis simply transmutes the blocking problem into a rule-ordering problem.

However Baker was not alone in thinking that the direct competition between two equivalent forms might somehow provide an answer to Gold's problem. Over the years, there have been many attempts to view blocking as a key to language learning.

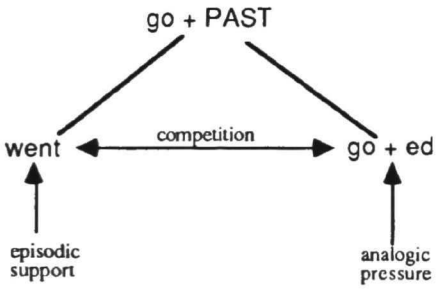
1. MacWhinney (1978) described the relation between "went" and "*goed" as a competition between rote and combination.
2. Anderson (1977; 1976) analyzed blocking relations in terms of a "semantic mapping" grammar that required a unique meaning for every form. Such grammars subscribe to Bolinger's (1977) principle that "when I say something different, I mean something different by it."
3. Wexler and Culicover (1980) introduced a Uniqueness Principle which also required a unique base structure representation for every surface structure form.
4. Pinker (1984) noted that blocking could be viewed as an instantiation of this Uniqueness Principle.
5. Clark (1983; 1987) emphasized the pragmatic basis of blocking in terms of her Principle of Contrast.
6. Markman (1989) proposed a blocking principle for lexical relations which she called Mutual Exclusivity.
7. Keil (1979) analyzed a particular form of cross-category blocking which he called the M-constraint.

Although these various accounts have intuitive appeal as mechanisms in language learning, the relation between blocking and Gold's problem has seldom been overtly analyzed. When it has been analyzed, the results fail to show any special power for blocking relations. For example, Kelly (19xx) and Anderson (1976) showed that, by itself, the Uniqueness constraint is not enough to block Gold's results. It appears that we need to look beyond mere blocking for a better understanding of the ways the learner deals with Gold's problem.

Competition

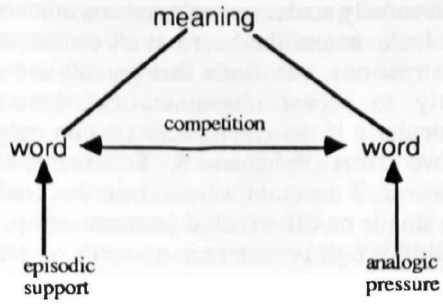
One can usefully reinterpret blocking phenomena in terms of a more general process called **competition** (MacWhinney, 1987, 1988, 1989). Competition applies not only to benign overgeneralizations, but also to virtually all lexical overgeneralizations produced by language learners. The crucial feature that distinguishes the Competition Model account from the blocking accounts of overgeneralization is its emphasis on the distinction between comprehension and production.

The Competition Model views all overgeneralizations as subject to three types of pressures. The first is the underlying analogic pressure that leads to the overgeneralization in the first place. This force impinges only on the production of the overgeneralization, not on its stored auditory form. The second pressure is the competition of the overgeneralized form with forms with confirmed auditory representations. The third pressure is the growth in the auditory representation of a correct form that accrues over time, as it is repeatedly encountered in the input data. Consider the case of "*goed" and "went" viewed diagrammatically. Here we have a overgeneralized form supported by analogy competing against a weak rote form supported by auditory memory:



As the strength of the auditory form for "went" grows, it competes more and more strongly against "*goed" until the error is almost entirely eliminated.

This is a specific instantiation of a more general schema for competition that looks like this:



Analogic Pressure

The competition between two candidate forms is governed by the strength of their episodic auditory representations. In the case of the competition between "*goed" and "went", the overgeneralized form has little episodic auditory strength, since it is heard seldom if at all in the input. Although "*goed" lacks auditory support, it has strong support from the general pattern for past tense formation (MacWhinney & Leinbach, 1991). In the Competition Model account (MacWhinney, 1978; 1988), it is analogic pressure that stimulates the overgeneralization and episodic auditory encoding that reigns it back in.

The analogic pressure hypothesized in this account has been described in detail in several connectionist models of morphophonological learning. Of these, the models that most closely implement the type of competition being described here are the models of MacWhinney and Leinbach (1991) for English and MacWhinney, Leinbach, Taraban, and McDonald (1989) for German. In these models, there is a general pressure for regularization according to the general pattern that produces forms such as "*goed" and "*ranned". In addition, there are weaker gang effects that lead to overgeneralizations such as "*stang" for the past tense of "sting".

Solving the LPLA

For these simple cases, the Competition Model solution to the LPLA is fairly simple. Recovery from overgeneralizations like "*goed" is guaranteed by the interaction between the three forces in the model: competition, analogic pressure, and episodic support. If analogic support is missing, no overgeneralizations will be produced. If analogic support is present, overgeneralizations will be produced. However, these forms will then enter into a (possibly

prolonged) competition with the correct irregular forms. Once the strength of the correct forms reaches a certain level, they will consistently dominate over the incorrect overgeneralizations and the child will no longer overgeneralize. Thus, there is no LPLA, at least for these simple cases.

Indirect Negative Evidence

The view of positive exemplars as providing negative evidence is akin to Chomsky's notion of "implicit negative evidence". According to Braine's (1989) elaboration of Chomsky's account, if the learner knows that "go" is a fairly frequent form, she can use the relative frequency of pairs such as "kick-kicked" or "cook-cooked" to infer that "*goed" is ungrammatical. Indirect negative evidence does not need to refer to the the relation between "*goed" and "went". However, like competition, it does need to keep good track of the strength of auditory forms. In addition, it needs to calculate the expected frequency of "*goed" based on other members of the paradigm. Competition, on the other hand, must rely directly on the increasing strength of "went". In this sense, the competition account relies more on blocking, whereas indirect negative evidence relies more on paradigmatic inference. Both accounts rely on the tracking of auditory strength. Although competition has more support in the empirical literature, both mechanisms are potentially useful in addressing the LPLA.

Both competition and the construction of negative evidence from positive evidence complicate the Gold framework by allowing more complex forms of negative evidence. It is hard to see how there is anything in Gold's framework that mitigates against these complications. However, in order to avoid confusion on these issues, it may be best to think of both competition and indirect negative evidence as providing frameworks through which learners construct "complex negative evidence". This evidence is not the direct provision of asterisks envisioned in the Gold paradigm. The parent does not raise an eyebrow or yell whenever the child makes a mistake. Instead, the child tracks more subtle distributional properties in the input which, together, can then be used to infer ungrammaticality. It the end, it is still the ungrammaticality of a form which guarantees recovery from overgeneralization and, in this sense, Gold's framework is maintained. However, what disappears is the LPLA, since we now see that there is sufficient evidence in the input to guarantee recovery from overgeneralization.

A reasonable reaction to this rather simple "*goed-went" demonstration is that the Competition Model account only works for trivial or "benign" cases and will surely fail on anything more challenging. However, it is possible to show that more complicated examples are explained by the same interplay of forces found in these simple examples.

Morphological Competition

Let us consider some more challenging cases of morphological overgeneralization, such as the overgeneralization "*unsqueeze" cited by Bowerman (1987) as particularly problematic for a Competition Model account. If the word "*unsqueeze" is being used to refer to the voluntary opening of a clenched first, likely competitors include "release" or "let go." Thus competition can work directly to weaken the overgeneralization. At the same time, the indirect negative evidence tracker can note that, although "squeeze" occurs frequently in the input, "*unsqueeze" does not.

It would be misleading to suggest that forms like "*unsqueeze" are handled only on the basis of raw occurrence data. There are also interesting semantic cues that support certain uses of this form while disallowing others. In particular, inanimate objects such as rubber balls and sponges cannot be "*unsqueezed" in the same way that they can be squeezed and that, in this sense, squeezing is only reversible for animate agents acting on themselves.

Similar forces operate on the competition between "*unapprove" and "disapprove". As the meaning of "*unapprove" moves farther away from a direct competition with "disapprove", it becomes correspondingly more acceptable. For example, we could imagine that a mortgage loan application which has been initially approved can then be subsequently "unapproved". At that point, we would still not have heard "unapproved" actually supported by input data, but there would be less direct competition with "disapprove". Forces that minimize the competition between meanings can help an overgeneralization survive long enough for it to begin to carve out its own "ecological niche" (MacWhinney, 1989).

Lexical Competitions

The same logic that can be used to account for recovery from morphological overgeneralizations

can be used to account for recovery from lexical overgeneralizations. For example, a child may overgeneralize the word "kitty" to refer to tigers and lions. The child will eventually learn the correct names for these animals and restrict the overgeneralized form. The same three forces are at work here: analogic pressure, competition, and episodic encoding. Although the child has never actually seen a "kitty" that looks like a tiger, there are enough shared features to license the generalization. If the parent supplies the name "tiger", there is a new episodic encoding which then begins to compete with the analogic pressure. If no new name is supplied, the child may still begin to accumulate some negative evidence, noting that this particular use of "kitty" is not yet confirmed in the input.

Competition tends to mimic the effects of the "mutual exclusivity bias" -- a tendency to treat each object as having only one name. However, since competition is implemented probabilistically through fuzzy logic (Massaro, 1987) or connectionist nets, it only imposes a bias, rather than a fixed constraint. The probabilistic basis for competition allows the child to deal with hierarchical category structure without having to enforce any major conceptual reorganization (Carey, 1985). Competition may initially lead a child to avoid referring to a "robin" as a "bird", since the form "robin" would be a direct match. However, there are a variety of circumstances in which "bird" no longer directly competes with "robin". These include reference to a collection of different types of birds that may include robins, reference to an object that cannot be clealry identified as a robin, or anaphoric reference to an item that was earlier mentioned as a "robin."

Syntactic Frame Competition

A critic could argue that the solution of Gold's problem for morphological structures tells us fairly little about the crucial syntactic overgeneralizations that are central to linguistic theory. According to this critic, showing that these "benign" violations are not problematic is uninteresting, since it is the difficult syntactic overgeneralizations that motivate the LPLA and its use as a support for nativist theory.

We can agree that, if the Competition Model account is to be at all general, it will also need to be be able to explain recovery from syntactic overgeneralization. Interestingly, the extension of the model to syntactic competition is fairly straightforward. The framework for this extension has been laid out in earlier papers (MacWhinney, 1982, 1987, 1988, 1989) which have described sentence processing in terms of lexical frame unification, much as in Kempen and Vosse (1989) or as in work in LFG. These frames specify slots or valency roles for one, two, or three arguments. For example, the verb "give" is a three-place predicate specifying slots for a giver, a recipient, and a transferred object.

Overgeneralizations in syntax arise when a valency pattern common to a large group of verbs is incorrectly overextended to a new verb. This type of overextension is described in connectionist networks such as those of McClelland and Kawamoto (1986) or Miikkulainen and Dyer (1990). These networks demonstrate the same gang effects and generalizations found in networks for morphological forms (Plunkett & Marchman, 1991) and spelling correspondences (Seidenberg, 1992; Seidenberg & McClelland, 1989). If a word shares a variety of semantic features with a group of other words, it will be treated syntactically as a member of the group. In the specific instantiation of this model explored in MacWhinney (1993), group membership can lead to the formation of new lexical structures, including new argument structures.

Consider the example of overgeneralizations of dative movement. Verbs like "give", "send", and "ship" all share a set of semantic features involving the transfer of an object through some physical medium. In this regard, they are quite close to a verb like "deliver" and the three-argument group exerts strong analogic pressure on the verb "deliver". However, for reasons that are still subjects of discussion (Gropen, Pinker, Hollander, Goldberg, & Wilson, 1989; Pinker, 1989), dative movement only applies to the frequent, monosyllabic transfer verbs and not to "deliver." When a child overgeneralizes and says "Tom delivered the library the book", she is illustrating the strength of the underlying analogic pressure of the group of transfer verbs that permit dative movement. In effect, the child has created a new argument frame for the verb "deliver". The first argument frame only specifies two arguments -- a subject or "giver" and an object or "thing transferred". The new lexical entry specifies three arguments. These two homophonous entries for "deliver" are now in competition, just as "*goed" and "went" were in competition. Like the entry for "*goed", the three-place entry for "deliver" has good analogic support, but no support from episodic encoding derived from the input. Over time, it loses in its competition with the two-argument form of "deliver" and its progressive weakening along with strengthening of the competing form leads

to recovery from overgeneralization. Thus, the analysis of recovery from "Tom delivered the library the book" is identical to the analysis of the "benign" case of recovery from "*goed".

Cue construction

Recovery from causative overgeneralizations such as "*I untied my shoes loose" works in a similar fashion. This particular extension receives analogic support from verbs like "shake" or "kick" which permit "I shook my shoes loose" or "I kicked my shoes loose". The restriction of this pattern to non-reversative verbs is semantically restricted, but it appears that the child is not initially tuned in to the fine details of these semantic classifications. Bowerman (1982; 1988) has suggested that the process of recovery from overgeneralization may focus the child's attention on extracting new features to block overgeneralization and this seems reasonable. Implementation of a connectionist model to simulate such error-based cue construction would be an interesting task.

Other causative overgeneralization types can be given similar accounts. For example, an error such as "*The gardener watered the tulips flat" can be attributed to a derivational pattern which yields three-argument verbs from "hammer" or "rake", as in "The gardener raked the grass flat." Source-goal overgeneralization can also fit into this framework. Consider "*The maid poured the tub with water" instead of "The maid poured water into the tub" and "*The maid filled water into the tub" instead of "The maid poured water into the tub". In each case, the analogic pressure from one group of words leads to the establishment of a case frame that is incorrect for a particular verb.

A more complex type of error involves not the establishment of a new three-place verb, but the use of a three-place verb in a non-licensed context. Bowerman (personal communication) provides the example "*The customers drove the taxi driver sad." Since one can say "The customers drove the taxi driver crazy", the error involves the exact shape of the resultative adjective. A connectionist model of the creation of a three-argument case frame for "drive" would require not only that certain verbs should license a third possible argument, but also what the exact semantic shape of that argument can be. In the case of the standard pattern for verbs like "drive", the resultant state must be terminative, rather than transient. To express this within the Competition Model context, we would need to have a competition between a confirmed three-argument form for "drive" and a looser overgeneral form based only on analogic pressure. A similar competition account can be used to account for recovery from an error such as "*The workers unloaded the truck empty" which contrasts with "The workers loaded the truck full". In both of these cases, analogic pressure seems fairly weak and examples of such errors are extremely rare in the language learning literature.

The actual modelling of these competitions in a network architecture will require detailed lexical work and extensive corpus analysis. Although there are a large variety of connectionist architectures that can implement the underlying analogic pressures involved, not all of them will deal cleanly with the episodic encoding required for governing lexical competition. A sketch of the types of models that will be required is given in MacWhiney (1993).

Monitoring

The account we have been examining holds that, over time, correct forms gain strength from encounters with positive exemplars and that this increasing stength leads them to drive out incorrect forms. In the terms of the notion of identification in the limit, this process is enough to guarantee the learnability of language. But the actual process of language learning includes a great deal more proactive processing that envisioned in this simple account. Consider a standard self-correction such as: I gived, uh, gave my friend a peach. Here the correct form "gave" is activated in real time just after the production of the overgeneralization. MacWhinney (1978) treated this type of self-correction as involving "expressive monitoring" in which the child listens to her own output, compares the correct weak rote form with the incorrect overgeneralization, and attempts to block the output of the incorrect form. One possible outcome of expressive monitoring is the strengthening of the weak rote form and weakening of the analogic forms. Exactly how this is implemented will vary from model to model

In general, retraced false starts move from incorrect forms to correct forms, indicating that the incorrect forms are produced quickly, whereas the incorrect rote forms take time to activate. Kawamoto (1992) has shown how a fully recurrent connectionist network can simulate exactly these timing asymmetries between analogic and rote retrieval. For example, Kawamoto's model captures the experimental finding that incorrect regularized pronunciations of "pint" to rhyme with "hint" are produced faster than correct irregular pronunciations.

Although expressive monitoring is a potential locus for much learning, an even more powerful learning mechanism is what MacWhinney (1978) called "receptive monitoring". If the child shadows input structures closely, he will be able to pick up many discrepancies between his own productive system and the forms he hears. Berwick (1987) found that a great deal of syntactic learning can be driven by the attempt to extract meaning during comprehension. Whenever the child cannot parse an input sentence, the failure to parse can be used as a means of expanding the grammar. The kind of analysis through synthesis that occurs in some parsing systems can make powerful use of positive instances to establish new syntactic frames. Receptive monitoring can also be used to recover from overgeneralization. The child may monitor the form "went" in the input and attempt to use his own grammar to match that input. If the result of the receptive monitoring is "*goed", the child can use the mismatch to reset the weights in the analogic system to avoid future overgeneralizations.

Conservatism

Many children are able to avoid falling into the trap of overgeneralization by using linguistic forms cautiously and conservatively. If a child avoids using a verb with dative movement until that verb is detected in a sentence with dative movement, dative movement overgeneralization will never occur. In general, conservative learners can learn without negative evidence, because they never make errors. Baker (1979), Fodor and Crain (1987), Maratsos, Kuczaj, Fox, and Chalkley (1979) and others have emphasized the extent to which syntactic learning can proceed on conservatively, often avoiding the need for negative evidence. O'Grady (O'Grady, 1987) and Parker (1989) have shown that conservatism can be used to account for learner acquisition of the sentence patterns that are often suggested as motivating the subjacency constraint and its related parameter. Of course, not all children are born conservative and overgeneralizations do occur, so our theories must still provide some means of explaining recovery from overgeneralization.

Rethinking the LPLA

The analysis we have presented views language learning as richly overdetermined. The forces that can work together to control overgeneralization include:

1. competition,
2. conservatism,
3. computation of indirect negative evidence,
4. cue construction,
5. expressive monitoring,
6. receptive monitoring,
7. working with complex negative evidence, and
8. occasional use of overt negative evidence.

Together, this rich armory of learning mechanisms indicates that the logical problem of language learning is easily solved, and that there is really no logical problem of language acquisition at all.

We need now to move beyond the LPLA to begin to look at errors and overgeneralizations in terms of the interplay of underlying analogic pressures and the social and cognitive forces that bring these pressures under control. In particular, we need to realize that parents are using the language learning situation not as a way of correcting their children, but as a way of developing mutual understandings (MacWhinney, 1985). When the parent knows what the child is talking about and when the child knows what the parent is talking about, positive instances are reinforced and, ultimately, it is positive instances, not negative instances that drive language learning and the correction of error.

References

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

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

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry, 10*, 533-581.

Baker, C. L., & McCarthy, J. J. (Ed.). (1981). *The logical problem of language acquisition*. Cambridge: MIT Press.

Berwick, R. (1987). Parsability and learnability. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition*. Hillsdale, N.J.: Lawrence Erlbaum.

Bohannon, N., MacWhinney, B., & Snow, C. (1990). No negative evidence revisited: Beyond learnability or who has to prove what to whom. *Developmental Psychology, 26*, 221-226.

Bohannon, N., & Stanowicz, L. (1988). The issue of negative evidence: Adult responses to children's language errors. *Developmental Psychology, 24*, 684-689.

Bolinger, D. (1977). *Meaning and form*. London: Longmans.

Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In E. Wanner, & L. Gleitman (Ed.), *Language acquisition: The state of the art*. New York: Cambridge University Press.

Bowerman, M. (1987). Commentary. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, N.J.: Lawrence Erlbaum.

Bowerman, M. (1988). The "no negative evidence" problem. In J. Hawkins (Ed.), *Explaining language universals*. London: Blackwell.

Braine, M. D. S. (1989). Modeling the acquisition of linguistic structure. In Y. Levy, I. Schlesinger, & M. Braine (Ed.), *Categories and processes in language acquisition*. Hillsdale, N.J.: Erlbaum.

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: M.I.T. Press.

Chomsky, N. (1963). Formal properties of grammars. In R. B. R. Luce, & E. Galanter (Ed.), *Handbook of mathematical psychology*. New York, Wiley Inc.

Chomsky, N. (1981). *Lectures on government and binding*. Cinnaminson, N.J.: Foris.

Clark, E. (1983). Meanings and concepts. In J. Flavell, & E. Markman (Ed.), *Handbook of Child Psychology: Vol. 3. Cognitive Development*. New York: Wiley.

Clark, E. (1987). The Principle of Contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition*. Hillsdale, N.J.: Erlbaum.

Demetras, M., Post, K., & Snow, C. (1986). Feedback to first-language learners. *Journal of Child Language, 13*,

Farrar, J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology, 28*, 90-98.

Fodor, J., & Crain, S. (1987). Simplicity and generality of rules in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition*. Hillsdale, N.J.: Lawrence Erlbaum.

Gass, S., & Schachter, J. (1989). *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press.

Gold, E. (1967). Language identification in the limit. *Information and Control, 10*, 447-474.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., & Wilson, R. (1989). The learnability and acquisition of the dative alternation in English. *Language, 65*, 203-257.

Hirsh-Pasek, K., Trieman, R., & Schneiderman, M. (1984). Brown and Hanlon revisited: mother sensitivity to grammatical form. *Journal of Child Language, 11*, 81-88.

Horning, J. J. (1969). *A study of grammatical inference*. Stanford University, Computer Science Department.

Kawamoto, A. (1992). Non-linear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account.

Keil, F. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.

Kempen, G., & Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: A cognitive architecture based on activation decay and simulated annealing. *Connection Science, 1*, 273-290.

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, N. J.: Lawrence Erlbaum.

MacWhinney, B. (1985). Grammatical devices for sharing points. In R. Schiefelbusch (Ed.), *Communicative competence: Acquisition and intervention*. Baltimore: University Park Press.

MacWhinney, B. (1987). Toward a psycholinguistically plausible parser. In S. Thomason (Ed.), *Proceedings of the Eastern States Conference on Linguistics*. Columbus, Ohio: Ohio State University.

MacWhinney, B. (1988). Competition and teachability. In R. Schiefelbusch, & M. Rice (Ed.), *The teachability of language*. New York: Cambridge University Press.

MacWhinney, B. (1989). Competition and lexical categorization. In R. Corrigan, F. Eckman, & M. Noonan (Ed.), *Linguistic categorization*. New York: Benjamins.

MacWhinney, B. (1993). The dinosaurs and the ring. In R. Corrigan Lima, S., Noonan, M. (Ed.), *The reality of linguistic rules*. San Diego: Academic.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 29*, 121-157.

MacWhinney, B., Leinbach, J., Taraban, R., & McDonald, J. (1989). Language learning: Cues or rules? *Journal of Memory and Language, 28*, 255-277.

Maratsos, M., Kuczaj, S. A., Fox, D. E., & Chalkley, M. A. (1979). Some empirical studies in the acquisition of transformational relations: Passives, negatives, and the past tense. In W. A. Collins (Ed.), *Children's language and communication.* Hillsdale, N.J.: Lawrence Erlbaum.

Marcus, G. (1993). Negative evidence in language acquisition. *Cognition, (in press),*

Marcus, G., Ullman, M., Pinker, S., Hollander, M., Rosen, T., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57,*

Markman, E. (1989). *Categorization and naming in children: problems of induction.* Cambrdige, MA: MIT Press.

Massaro, D. (1987). *Speech perception by ear and eye.* Hillsdale, NJ: Lawrence Erlbaum.

Matthews, R., & Demopoulos, W. (1989). *Learnability and linguistic theory.* Dordrecht: Kluwer.

McClelland, J., & Kawamoto, A. (1986). Mechanisms of sentence processing: Assigning roles to constituents. In J. McClelland, & D. Rumelhart (Ed.), *Parallel Distributed Processing.* Cambridge, MA: MIT Press.

McNeill, D. (1966). Developmental psycholinguistics. In F. Smith, & G. A. Miller (Ed.), *The genesis of language: A psycholinguistic approach.* Cambridge, Mass.: MIT Press.

Miikkulainen, R. (1990). A distributed feature map model of the lexicon. In (Ed.), *Proceedings of the 12th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Morgan, J., & Travis, L. (1989). Limits on negative information in language input. *Journal of Child Language, 16,* 531-552.

O'Grady, W. (1987). *Principles of grammar and learning.* Chicago: Chicago University Press.

Parker, K. (1989). Learnability theory and the acquisition of syntax. *University of Hawaii Working Papers in ESL, 8,* 49-78.

Penner, S. G. (1987). Parental responses to grammatical and ungrammatical child utterances. *Child Development, 58,* 376-384.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, Mass: Harvard University Press.

Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure.* Cambridge: MIT Press.

Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition, 38,* 43-102.

Seidenberg, M. (1992). Connectionism without tears. In S. Davis (Ed.), *Connectionism: Theory and practice.*

Seidenberg, M., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.

Sokolov, J. (1990). A distributional analysis of the fine-tuning hypothesis. *Developmental Psychology, submitted,*

Weissenborn, J., Goodluck, H., & Roeper, T. (1992). *Theoretical issues in language acquisition: Continuity and change in development.* Hillsdale, NJ: Lawrence Erlbaum Associates.

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