ON RULES AND EXCEPTIONS: AN INVESTIGATION OF INFLECTIONAL MORPHOLOGY

by

Gary Fred Marcus

SUBMITTED TO THE DEPARTMENT OF BRAIN AND COGNITIVE SCIENCES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1993

Signature of Author ______________________________

Department of Brain and Cognitive Sciences
June 3, 1993

Certified by ______________________________

Professor Steven Pinker
Thesis Supervisor

Accepted by ______________________________

Professor Emilio Bizzi
Chairman, Department of Brain and Cognitive Sciences

ON RULES AND EXCEPTIONS: AN INVESTIGATION OF INFLECTIONAL MORPHOLOGY

by

Gary Fred Marcus

Submitted to the Department of Brain and Cognitive Sciences on June 3, 1993 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Abstract

This thesis examines two theories of the acquisition and representation of inflection (e.g. English past tense formation), focusing on children's overregularization errors such as *goed* and *wented.* On the "rule" view, suggested by Pinker and Prince (1988), regular forms (e.g. *walk-walked*) are created by the on-line application of symbolic, default rule (add *-ed* to form the past tense), while irregular forms are retrieved from an associative memory. The alternative "connectionist" view, suggested by Rumelhart and McClelland, is that systems of rules and exceptions could be represented by single uniform associative connectionist network which contains no rules. These two views are considered with respect to four systems of inflection: the formation of plural and past tense forms in English, and plural and participle forms in German.

Parts 2 and 3 focus on children's overregularization errors, such as *foots* and *mans*. On the symbolic view, such errors are caused by the application of the default rule whenver an irregular cannot be retrieved from memory. On the connectionist view, overregularizations are caused when irregular words are drawn by an analogical process to regular words. A study of 11,521 irregular past tense utterances in the spontaneous speech of 83 children. The major results were as follows: (1) Overregularization errors are relatively rare (median 2.5% of irregular past tense forms), suggesting that there is no qualitative defect in children's grammars that must be unlearned. (2) Overregularization occurs at a roughly constant low rate from the two's into the school-age years, affecting most irregular verbs. (3) Though overregularization errors never predominate, one aspect of their purported U-shaped development was confirmed quantitatively: an extended period of correct performance precedes the first error. (4) Overregularization does not correlate with increases in the number or proportion of regular verbs in parental speech, children's speech, or children's vocabularies. Thus the traditional account in which memory operates before rules cannot be replaced by a connectionist alternative in which a single network displays rotelike or rulelike behavior in response to changes in input statistics. (5) Overregularizations first appear when children begin to mark regular verbs for tense reliably (i.e., when they stop saying *Yesterday I walk*). (6) The more often a parent uses an irregular form, the less often the child overregularizes it. (7) Verbs are protected from overregularization by similar-sounding irregulars, but are not attracted to overregularization by similar-sounding regulars, suggesting that irregular patterns are stored in an associative memory with connectionist properties, but regulars are not.

A study of English pluralization confirms these basic findings. The data are from the speech of 10 English speaking children taken from the CHILDES database (MacWhinney and Snow 1985, 1990). The overall rate of noun overregularization is low, mean = 8.5%, demonstrating that children prefer correct to overregularized forms. Children's rates of noun overregularization are not significantly different from their rates of verb overregularization, and noun plurals, like verb past tenses, follow a u-shaped developmental curve in which correct irregulars precede the first overregularized forms. These facts suggest that plural and past tense overregularizations are caused by similar underlying processes. The results pose challenges to connectionist models, but are consistent with Marcus, et al.'s (1992) blocking-and-retrieval-failure model in which regulars are generated by a default rule while irregulars are retrieved from the lexicon.

Contrary to connectionist accounts, these effects are not due to regular words being in the majority. The German participle *-t* and plural *-s* apply to minorities of words. Two experiments eliciting ratings of novel German words show that the affixes behave like their English counterparts, as defaults. Thus default suffixation is not due to numerous regular words reinforcing a pattern in associative memory, but to a memory-independent, symbol-concatenating mental operation.

Thesis Supervisor: Professor Steven Pinker
Title: Professor of Brain and Cognitive Sciences

Acknowledgments

I couldn't have had a better role model than my advisor and collaborator, Steve Pinker. Even after working with Steve for four years, I am amazed daily by his clarity and brilliance as a researcher, speaker, writer, teacher, and advisor. The care, committment, creativity, and rigor that he brings to science is inspiring. It's an understatement to say that this thesis could not have been written without his help.

Indeed, this work has profited greatly from a long list of coauthors: Ursula Brinkmann, Harald Clahsen, Michelle Hollander, Steve Pinker, T. John Rosen, Michael Ullman, Richard Wiese, Andreas Woest, and Fei Xu. Thanks also to Hartley Kuhn, Elena Augustino, and Elliza Mcgrand who helped collect some of the data in Part 2.

For many useful discussions on a broad range of topics in cognitive science, I thank my thesis committee Steve Pinker, Susan Carey, Ken Wexler, Morris Halle, and, unofficially, Harald Clahsen. I have also gained a great deal from other faculty, especially Alec Marantz, David Pesetsky, and Molly Potter.

My family, Mom, Dad, Linda, and Julie provided love, support and encouragement. I'd especially like to thank my dad for always working hard to understand and challenge my work. Julie (age 4;10) for her part overregularized, but not too often.

Thanks to Neal Pearlmutter, for endless conversation, and for never asking me to clean up our office; to Ed Stein for making sure my work stood on firm philosophical ground; to David Poeppel for indefatigably searching his mental lexicon and answering all my questions about German; and to two friends, Zach and Justin, who aren't cognitive scientists but who listened anyway.

For comments, criticism, and encouragement, I thank my fellow graduate students including Sergey Avrutin, Paul Bloom, Kevin Broiher, Laura Conway, Tony Harris, Suzie Johnson, John Kim, Eric Loeb, Neal Pearlmutter, Colin Philips, David Poeppel, Sandeep Prasada, Ed Stein, Josh Stern, William Snyder, Michael Ullman, and Fei Xu.

Thanks also go to Jan Ellertsen for helping me navigate through the bureaucracy, to Pat Claffey, the nicest and most interested librarian anyone could hope for, and to Robin Nahmias and Ellie Bonsaint for administrative support.

Throughout my graduate career, Fei has been a true friend, invaluable colleague, and more. Without Fei's constant support, I couldn't have written this thesis, which I dedicate to her.

This research was supported a NDSE Graduate Fellowship to me, a grant of the German Science Foundation to Harald Clahsen (grant-No Cl97/5-1, Wu86/9.1-9.4), NIH Grant HD 18381 and NSF Grant BNS 91-09766 to Steven Pinker, and grants to the Department of Brain and Cognitive Sciences from NIMH (training grant T32 MH18823) and the McDonnell-Pew Program in Cognitive Neuroscience, and a grant from the American

Council of Learned Societies and the Deutscher Akademischer Austauschdienst German-American Collaborative Research awarded jointly to Clahsen and Pinker.

Preface

This thesis is divided into four Parts that were written to stand separately, and hence refer to each other as separate entities.

Part I is concerned with the question of whether children require "negative evidence" (i.e., parental feedback about which strings of words are not grammatical sentences) to eliminate their ungrammatical utterances. Lacking negative evidence, a child would require internal mechanisms to unlearn grammatical errors. Several recent studies claim that parents tend to react differently to their children depending on whether children speak correctly or make grammatical errors. For example, parents are more likely to repeat grammatical sentences verbatim than ungrammatical sentences. These patterns are argued to be provided to the child as corrective feedback.

Part I argues that this sort of feedback is too statistically unreliable to be able to have any effect. Moreover, this feedback is too variable, both within individual studies and across cultures, to be necessary for language acquisition. Finally, I show that the claim that parents respond differently according to whether children speak grammatically was based on a methodological artifact. Therefore, corrective feedback is unlikely to be necessary for language acquisition; instead language acquisition must depend on mechanisms internal to the child.

Part I thereby motivates a central question in language acquisition: lacking corrective feedback, how do children unlearn errors such as past tense overregularizations (e.g. *breaked*? Parts II and III present a theory of how children unlearn such errors based on quantitative analysis of spontaneous speech.

In Part II, I analyzed 11,521 past tense utterances from the spontaneous speech of 83 children. I discovered that -- contrary to popular opinion reported everywhere from

Newsweek to textbooks to articles in the primary literature -- there is no stage in which children completely replace correct forms with overregularizations. Instead, children overregularize rarely, in only about 4% of their opportunities. Moreover, by analyzing the time course of overregularization we were able to confirm quantitatively for the first time one aspect of the u-shaped developmental sequence: a period of correct performance does precede the earliest overregularizations. Furthermore, we tested a prediction of connectionist models which predict that a child displays rotelike or rulelike behavior in response to changes in input statistics. This prediction was incorrect: overregularization does not correlate with increases in the number or proportion of regular verbs in parental speech, children's speech, or children's vocabularies. Instead, children begin to overregularize when they begin to reliably mark the past tense of regular verbs, suggesting that overregularization is tied to the acquisition of the regular past tense rule. A wide range of data can be accounted for by a simple model: Irregular forms are retrieved from memory and block the application of the default regular rule (add -ed); if a child fails to retrieve the irregular form the regular rule applies by default and the child produces an overregularization.

Part III presents a quantitative study of a closely related error, noun plural overregularizations (e.g., *foots* and *mans*). The data are from the speech of 10 English speaking children taken from the CHILDES database (MacWhinney and Snow 1985, 1990). The overall rate of noun overregularization is low, mean = 8.5%, demonstrating that children prefer correct to overregularized forms. Children's rates of noun overregularization are not significantly different from their rates of verb overregularization, and noun plurals, like verb past tenses, follow a u-shaped developmental curve in which correct irregulars precede the first overregularized forms. These facts suggest that plural and past tense overregularizations are caused by similar underlying processes. The results pose challenges to connectionist models, but are consistent with the blocking-and-retrieval-failure model, presented in part II, in which regulars are generated by a default rule while irregulars are retrieved from the lexicon.

Part IV uses cross-linguistic evidence to tease apart connectionist and symbolic models of language. In English, most novel verbs form their past tense with -ed, e.g., fax-faxed. In connectionist models, the verb fax would be inflected with -ed because -ed is the most frequent way of forming the past tense. In symbolic models such as the one argued for in Parts 2 and 3, the past tense of fax would not be found in the lexicon, and hence the default rule (add -ed), would apply, creating faxed. The operation of this default rule is independent of any given pattern's frequency. Although these models are quite different, they are difficult to distinguish in English, since the -ed pattern is so frequent. By picking a language, namely German, in which no single regular pattern is predominant, I was able to test whether rule use depends on high frequency. In experiments with adult speakers, I discovered that even though German -s plural applies to only a tiny minority of German nouns, it is used as the plural for unusual sounding nouns, names, borrowings, and other circumstances in which no plural is listed in the lexicon. Thus, default inflection does not depend on frequency, but rather on whether an inflected form is available from the lexicon, as predicted by symbolic but not connectionist models.

Part I previously appeared as Marcus, G. F. (1993). Negative evidence in language acquisition. Cognition, 46, 53-85.

Part II previously appeared as Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J. and F. Xu (1992). Overregularization in language acquisition. Monographs of the Society for Research in Child Development., 57 (4, Serial No. 228).

Part III has been submitted to a peer-review journal.

Part IV previously appeared as Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., Woest, A., and Pinker, S. (1993). German inflection: The exception that proves the rule. MIT Center for Cognitive Science Occasional Paper #47, Cambridge, MA.

Table of Contents

PART I

Negative Evidence in Language Acquisition

Introduction

A major challenge for theories of language acquisition is to explain how children recover from grammatical errors, such as (1) and (2):

(1) *I maked it with water. (Sarah 4;5; Brown, 1973)

(2) *And fill the little sugars up in the bowl. (Mark 4;7; Pinker, 1989)

Do children eliminate their grammatical errors solely on the basis of internal mechanisms, or do they require external feedback from their parents?

Brown and Hanlon (1970) carried out two analyses in the first systematic examination of whether parents provide feedback contingent on children's grammatical errors. In the first, they examined whether parents comprehend their children better if their children's utterances are grammatical. Each child question was classified as grammatical or ungrammatical; each parental reply was coded to indicate whether the parent understood or failed to understand the child's utterance. Replies indicating the lack of comprehension followed about as many grammatical child questions (42%) as ungrammatical ones (47%); similarly, replies indicating comprehension were equally likely following grammatical and ungrammatical speech (45% in both cases). Brown and Hanlon concluded that "In general, the results provide no support for the notion that there is a communication pressure favoring mature constructions" (p. 45).

In their second analysis, Brown and Hanlon coded whether parental replies indicated approval (e.g., *That's right* or *Yes*) or disapproval (e.g., *That's wrong* or *No*). Again, there was no relation between parental reply types and child grammaticality. Finding not "even a shred of evidence that approval and disapproval are contingent on syntactic correctness" (p. 47), Brown and Hanlon concluded that:

> While there are several bases for approval and disapproval, they are almost always semantic or phonological. Explicit approval or disapproval of either syntax or morphology is extremely rare in our records and so seems not to be the force propelling the child from immature to mature forms. (p. 48)

Based largely on their conclusions, much subsequent research in language acquisition has tried to solve the puzzle of how children acquire language solely from *positive evidence* (i.e., hearing sentences that belong to a language) and without *negative evidence* (i.e., information about which sentences do not belong to that language). (E.g., Braine, 1971; Wexler & Hamburger, 1973; Chomsky & Lasnik, 1977; Baker, 1979, 1981; Wexler & Culicover, 1980; Baltin, 1981; Grimshaw, 1981; Roeper, 1981; Bowerman, 1983, 1987, 1988; Mazurkewich & White, 1984; Pinker, 1984, 1989; Berwick, 1985; Lasnik, 1981, 1989; Crain and Fodor, 1987; Wexler & Manzini, 1987; Lightfoot, 1989; Matthews & Demopoulous, 1989; Gropen, Pinker, Hollander & Goldberg , 1991; Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992.) The "no negative evidence problem" -- that is, how children could learn language without negative evidence -- is often seen as necessarily tied to nativist explanations that posit the existence of internal and innate mechanisms. In fact, Braine (1971), the first to discuss the implications of the "no negative evidence" problem, used the problem to argue *against* Chomsky's (1965) nativist explanation of language acquisition. Arguing that Chomsky's hypothesis-testing proposal would require negative evidence, Braine presented anecdotal evidence suggesting that negative evidence was insufficiently frequent and perhaps ignored, hence refuting Chomsky's proposal. Without negative evidence, any model of language acquisition, nativist or not, must account for how the child can learn language from positive evidence alone. Suppose a child must learn that a certain sentence, B, is ungrammatical. The child might learn this in two ways.

First, something external to the child (i.e., explicit negative evidence) could tell the child that B is ungrammatical. If the parent does not provide an explicit denial of sentence B, the only alternative is that the parent says a sentence (or a set of sentences), A, and that the child has a mechanism which eliminates B when given A. That mechanism, innate or learned, must be internal to the child, regardless of whether it is, for example, a general pragmatic mechanism or a linguistically specific mechanism. The existence or non-existence of negative evidence allows us to determine whether internal mechanisms are needed. If negative evidence does not exist, the task of language acquisition researchers must be to discover which internal mechanisms do allow children to eliminate their errors.[1]

Recently, in what I will call Discourse Studies, some researchers have disputed the claim that there is no negative evidence (Hirsh-Pasek, Treiman, and Schneiderman, 1984; Demetras, Post, and Snow, 1986, Penner, 1987; Bohannon and Stanowicz, 1988; and Morgan and Travis, 1989). These studies, reporting evidence from the distribution of certain patterns of discourse between parents and children, argue that parents do provide implicit feedback to their children based on whether children speak grammatically.[2]

Each study used somewhat different reply categories; the Appendix reprints exact definitions. Table 1 provides rough definitions of each type of parental reply.

---

[1]"Indirect negative evidence" (e.g., Chomksy 1981) -- information about which sentences have not appeared in the input -- is sometimes offered as a third type of evidence available to the child. The use of indirect negative evidence would require an inference such as "If X appears in the input, assume that Y is not in the language, unless you hear Y in the input." Indirect negative evidence thus depends on a reanalysis of positive evidence based on mechanisms internal to the child, rather than input external to the child, hence for the purposes of this paper, I collapse it with positive evidence. See Pinker 1989:14-15 for further discussion.

[2]All of the studies that I will discuss contrast parental replies to grammatical sentences with parental replies to ungrammatical sentences. However, one study, Moerk (1991), presented examples only of how parents reply to ungrammatical sentences and failed to present examples of how parents reply to grammatical utterances. For example, in the following exchange, Moerk called the reply, What did I do? a "correction" (Eve: What did you doed? Mother: What did I do?), but he neglected to note that the identical parental reply, two samples earlier, is not a "correction." (Eve: Hurt my eye. Mother: Your eye? Eve: Yeah. Mother: What did I do?) (Eve 2;1). Moerk provided no objective definition of "corrections" and no explanation of how a child could recognize or use them. Applying this method, one could argue that the word red provides negative evidence: simply cull transcript examples in which the parent said red following a child's ungrammatical utterances and ignore the occasions in which the parent said red after the child's grammatical utterances. Without the crucial contrast between replies to grammatical and ungrammatical sentences, such studies are irrelevant to the question of whether parents provide negative evidence.

Table 1:
Definitions of Types of Parental Replies used in Discourse Studies

Explicit Approval	Parent says yes or un-huh or the like.
Non Sequiturs	Parent fails to understand child.
Repetitions	Parent repeats child utterance, (Hirsh-Pasek only et al.: with changes)
Imitations/Exact Reps.	Parent repeats child utterance, verbatim
Expansion	P. repeats child utterance, but makes grammatical changes or adds new material
Recasts	Parent repeats child utterances, but with minor grammatical changes.
Topic Extensions	Continued topic but not a Repetition or Expansion
Move-ons/No Responses	Parent moves conversation along
Clarification Questions	P. asks a question that requests child to repeat part of utterance.
Confirmation Questions	Parent asks a Yes/No question.

Some discourse patterns, such as Expansions (interchanges between a child and a parent in which a parent repeats the child's utterance with small changes), are claimed to be elicited more often by ungrammatical speech than by grammatical speech. Other discourse patterns such as Exact Repetitions (interchanges in which a parent exactly repeats a child utterance) are claimed to be elicited more often by grammatical speech than by ungrammatical speech.

In all cases, the differences in parental replies to grammatical versus ungrammatical sentences were statistical rather than categorical (i.e., all-or-none). I call this type of feedback *noisy feedback*, since parents provided each type of reply after both grammatical and ungrammatical sentences, albeit in different proportions.

In a recent discourse study, Bohannon and Stanowicz (1988) argue that discourse-based noisy feedback "may be considered superior to simple denials and as qualified to assume the role of 'negative evidence'" (p. 688), and concluded that "To the extent that current theories of language acquisition also ignore adults' tendency to provide feedback (i.e., negative and specific evidence), these theories will fail to accurately account for language acquisition" (p. 688).

These conclusions, if correct, would undermine much of the research based on the assumption that parents do not provide negative evidence. However, because the methodology and results of the Discourse Studies vary and often conflict with one another, and because there has been little discussion of how children could use noisy feedback to eliminate grammatical errors, the role and existence of noisy feedback remains controversial. This controversy is evident both empirically, in the conflicting results and conclusions of Bohannon and Stanowicz (1988) and Morgan and Travis (1989), and theoretically, in a recent debate between Gordon (1990) and Bohannon, MacWhinney, and Snow (1990).

The goal of this paper is to define and apply clear criteria for clarifying the role and existence of noisy feedback. These criteria will lead to the following conclusions:

-- Noisy feedback is too weak to be a plausible way of eliminating errors. Even under statistically optimal conditions, a child would have to repeat a given sentence verbatim more than 85 times to eliminate it from his or her grammar.

-- Noisy feedback is inconsistent across parents, declines or disappears with age, and is probably not provided for all types of errors and is thus unavailable for much of language acquisition. Furthermore, some apparent patterns of noisy feedback may be averaging artifacts that do not correspond to types of feedback given to any individual child.

-- Because parental reply categories are defined only with respect to the child's utterance, "correlations" between the two may be artifacts resulting from the definition of parental reply categories and of constant noncontingent properties of parental and child speech. For example, several studies show that verbatim repetitions follow more grammatical than ungrammatical sentences. But because parents virtually never speak ungrammatically, verbatim repetitions necessarily negatively correlate with children's ungrammatical utterances.

Any one of these three conclusions, the weakness, the inconsistency, or the inherently artifactual nature of noisy feedback, would by itself severely undermine, if not falsify, the position that parents provide feedback to their children and thus that such feedback is required for language acquisition. Of course, even if parents provided feedback, it might turn out empirically that children do not use it. For example, Zwicky (1970), in discussing his daughter's inflectional overregularizations (e.g., *goed*), points out that "six months of frequent corrections by her parents had no noticeable effect." Braine (1971) even more clearly illustrates that feedback may be ineffective.

> For experimental purposes, I have occasionally made an extensive effort to change the syntax of my two children through correction. One case was use by my two-and-a-half year-old daughter of *other one* as a noun modifier. Over a period of a few weeks I repeatedly but fruitlessly tried to persuade her to substitute *other+N* for *other one+N*. With different nouns on different occasions, the interchanges went somewhat as follows: "Want other one spoon, Daddy"—"You mean, you want THE OTHER SPOON"—"Yes, I want other one spoon, please, Daddy"—"Can you say 'the other spoon'?"—"Other . . . one . . . spoon"—"Say . . . 'other'"—"Other"—"Spoon"—""Spoon"—"Other . . . spoon" —" Other . . . spoon. Now give me other one spoon?" Further tuition is ruled out by her protest, vigorously supported by my wife. Examples indicating a similar difficulty in using negative information will probably be available to any reader who has tried to correct the grammar of a two-or-three-year old child. (p. 160-161)

Even when children understand that they are being corrected, they sometimes make incorrect generalizations, as MacNeill (1966, p. 69) shows:

> Child: Nobody don't like me. Mother: No, say "nobody likes me." Child: Nobody don't like me. [Eight repetitions of this dialogue follow.] Mother: No, now listen carefully, say "NOBODY LIKES ME." Child: Oh! Nobody don't likes me.

These examples suggest that it is the child's underlying linguistic system, rather than negative evidence, which forces children to change their grammars. Until the existence of reliable parental feedback is firmly established, however, it is premature to consider whether children actually use it.

Types of Parental Feedback and their role in Language Learnability

Positive evidence is simply the input, that is the sentences children hear. In contrast, *Negative evidence* is a parental behavior that provides information about which sentences are not in the language. Negative evidence does not tell a child which sentences are grammatical; rather, it indicates that the child has uttered an ungrammatical sentence. Moreover, negative evidence does not tell a child *why* a particular sentence is ungrammatical.

A child who does not speak can receive no negative evidence, aside from explicit metalinguistic statements (e.g., *don't say X*). Negative evidence must respond to a child's utterance; it can only arise if parents are in some way *sensitive* (though not necessarily consciously) to whether their children speak grammatically.

The parental behavior that provides negative evidence I will call the *reply type*. The reply type may take many forms: the parent might say *no*, provide a repetition, shrug, or even spank the child. An individual occurrence of a given reply type (e.g., a single utterance of the word *no*) is a *reply instance*.

To distinguish three kinds of negative evidence that are often collapsed, I use the terms *complete feedback*, *partial feedback*, and *noisy feedback*. *Complete feedback* is any corrective signal provided for all and only those sentences that are ungrammatical (c.f. Gold's (1967) term *informant presentation* in the formal mathematical literature on learnability). *Partial feedback* is any corrective signal provided following some ungrammatical sentences, but never provided following correct sentences. *Noisy feedback* is a corrective signal provided after some errors and after some correct sentences, but in different proportions. A single reply instance of noisy feedback, unlike instances of complete feedback and partial feedback, does not guarantee that a sentence is ungrammatical.

Verbal parental feedback, such as Repetitions or Expansions, inherently provides the child with positive evidence, as well as potentially providing negative evidence. Several studies have reported that Expansions may spur learning (though Morgan, Bonamo, and Travis, 1991, conclude the opposite). Crucially, however, even if children learn from Expansions, they might do so without using negative evidence. In principle, a child could learn whether some sentence is grammatical (a) only from positive evidence, which tells the child that the parental utterance is a possible way of saying what she wanted to say; (b) only from negative evidence, which tells the child that her utterance is not acceptable; or (c) both, which would tell the child both that her sentence was unacceptable and provide an acceptable alternative; each of these possibilities is plausible *a priori*. Consider this hypothetical dialogue:

(3) Child: I eated the food Parent: I ate the food.

This parental reply clearly provides the child with a piece of positive evidence: *I ate the food* is a grammatical sentence. Positive evidence alone does not tell the child whether *eated* and *ate* are stylistic variants or synonyms or whether *eated* is unacceptable, but might do so in combination with internal mechanisms. The parental reply could serve as negative evidence only if the child recognizes it as an exemplar of a Recast and has mechanisms that use such information.

Complete Feedback A child who received complete feedback might eventually converge on the correct language by eliminating all possible but ungrammatical sentences from her language (Gold, 1967; Pinker 1979). If complete feedback existed it would probably be easy to detect, since there would, by definition, be a perfect correlation between the reply type and the grammaticality of the child's utterance. Nobody, though, has ever provided any evidence that children receive complete feedback; every reply type studied in the discourse studies is not provided for all ungrammatical sentences, and is provided for some grammatical sentences. Nobody claims, for instance, that parents Expand children's utterances every time that a child errs but never when the child speaks grammatically. Thus children almost certainly do not receive complete feedback.

Partial Feedback Partial feedback, by definition, is not provided every time a child makes an error. There has been no explicit discussion of how children could use such feedback, although Gordon (1990) has suggested that "it may, in fact, be possible to show that learning can occur with inconsistent feedback . . . using statistical rather than absolute criteria to determine grammaticality" (p.219). It is not clear either whether children have mechanisms for using partial feedback, nor what those mechanisms might be.

In any case, partial feedback does not seem to exist. No reply type studied in the discourse studies is provided only for ungrammatical sentences but not for grammatical sentences, or vice-versa. The evidence from the discourse studies suggests that all patterns of discourse provide, at best, noisy feedback: every reply type is provided following both grammatical and ungrammatical speech.

Noisy feedback

Several studies claim to have documented patterns of noisy feedback, but there has been no explicit discussion of how children might use it. I will examine a concrete example, after defining four terms. The proportion of reply instances (e.g. Expansions) elicited by grammatical speech will be denoted by pr{gr}. The proportion of reply instances elicited by ungrammatical speech is, then, pr{ungr}. The number of times the child tests a given sentences is *n*; the observed proportion of reply instances is *p(obs)*. I will assume that the feedb .k is provided probabilistically and independently.

Suppose pr{ungr} = .20 and pr{gr} = .12 (these proportions are drawn from the distribution of noisy feedback reported by Hirsh-Pasek et al.) Suppose the child says a sentence once (i.e. $n = 1$). Depending on whether the parent provides a reply instance, *p(obs)* will be either 0 or 1. Suppose *p(obs)* is 1; given the small sample size, *p(obs)* could easily come from either distribution. The chance that *p(obs)* comes from pr{ungr} is .625 (since 20 out of every 32 reply instances would follow ungrammatical sentences). The chance that it comes from pr{gr} is .375 (= 1-.625). Thus if the child decides that the sentence is ungrammatical, there is a .375 chance that the child is wrong. Or if *p(obs)* is 0, there is a .52 (88/168) chance that the sentence is grammatical, and .48 chance that the sentence is ungrammatical. In either case, the child has some information but not enough.

The child's chances improve if the child increases the sample size, n, by repeating the sentence. Suppose after 100 repetitions of the sentence, *p(obs)* = .17. Then it is relatively likely that the sentence is from pr{ungr} but still reasonably likely that *p(obs)* comes from pr{gr}. If a child repeats the test sentence 1000 times, with *p(obs)* = .20, the child could decide with near certainty. In general, the more times the child repeats the same sentence, the more information the child gathers, and the easier it is for the child to determine the grammaticality of that sentence, with some degree of certainty; I will assume that the degree of certainty equals the rate of error.

Because the end point of language acquisition is the adult state, I assume that the acceptable rate of error, e , should reflect the rate at which adults make grammatical errors (or perhaps be lower, since adults often know when they have made errors). If e is too high, the child will fail to converge on the adult state of rarely making errors. The overall rate of adult past tense overregularization errors (e.g. *maked* or *Maded*) is around .00004 (Marcus et al. , 1992), while the overall rate of adult speech errors is about .001 (Stemberger, 1989). To be generous by an order of magnitude, suppose e is .01.[3]

---

[3]Actually, two error rates are involved here: the chance of incorrectly accepting a sentence as grammatical and the chance of incorrectly accepting a sentence as ungrammatical. Since most speakers agree both on which sentences are grammatical and on which sentences are ungrammatical, both rates should be very low . Lacking evidence to the contrary, I assume that these two error rates are equal.

In Figure 1, the left curve represents the sampling distribution of the proportion of parental reply instances out of child utterances provided following grammatical sentences, while the right curve represents the sampling distribution of the proportion of parental reply instances out of child utterances provided following ungrammatical sentences. The shapes of two curves depend on the proportions of feedback to grammatical and ungrammatical sentences and the number of times, n, that the child repeats the test sentence; here pr{gr} = .12, pr{ungr} = .20, $n = 10$.

Figure 1 Here

In order to decide whether a sentence is grammatical, assuming e= .01, a child must determine whether *p(obs)* is within the 99% confidence interval of the grammatical distribution, within the 99% confidence interval of the ungrammatical distribution, or within the overlap between the two confidence intervals, which I call *the region of ambiguity*, shown as the cross-hatched region in Figure 1. This region indicates possible values of p(obs) under which a child cannot make a decision within the specified risk of error.[4]

The only way the child can reduce the region of ambiguity is to repeat her test sentence again and again, n times. Figure 2 represents the smallest value of n in which the region of ambiguity is eliminated (for e = .01). Before being able to make a decision, the child would need to repeat the same construction 446 times.

Figure 2 Here

)

Of course, the minimum number of repetitions necessary to decide whether a

---

[4]Note that for positive evidence, there is virtually no region of ambiguity. If the parent says a sentence the child can be nearly certain that the sentence is grammatical. According to Newport, Gleitman, and Gleitman (1977), 99.44% of parental speech to children is grammatical. Similarly, if the parent provided complete feedback, there would also be no region of ambiguity. Upon hearing a reply instance, the child could immediately decide that a given sentence was ungrammatical.

sentence is grammatical depends on the proportions of feedback; these values vary across the Discourse Studies. Table 2 shows estimates of n for the clearest and easiest reply types the child could use. Weaker differences between parental replies to grammatical and ungrammatical sentences would require the child to repeat a test sentence even more often.

Table 2: Minimum number of times (*n*) a child would need to repeat a given sentence verbatim to decide whether it was grammatical, with chance of error, e, less than .01, calculated for the strongest reply types reported in the Discourse Studies. (N may be determined by finding the minimum number of repetitions of a single sentence such that the 99% confidence intervals of the two distributions (pr{gr}, pr{ungr}) do not overlap.)

Study	Reply Type	pr{ungr}	pr{gr}	n
Hirsh-Pasek....	Repetitions 2 yr. olds	20	12	446
	3-5 year olds no	differences	n/a	
Penner	Expansions: Group 1	18.3	4.6	104
	Expansions: Group 2	11.3	6.3	679
Boh,&Stan.	Total of All Reply Types	35	14	85
Morg.& Trav.	Expansions (A, E, & S)	11.3	2.7	169

Note: The data from Demetras et al. are excluded because most patterns of reply varied across parents and no inferential statistics (or sample sizes) were provided. Without such information we cannot infer the reliability of the data. The data for Morgan and Travis are an average across the patterns of parental expansions to Adam, Eve, and Sarah, from Table 4, Morgan & Travis 1989:545, because there is substantial variation between the pattern of expansions to the different children. Eve would only require 63 verbatim repetitions, but Sarah would require 300.

The lowest estimate of *n*, derived from data reported by Bohannon and Stanowicz (1988), is 85; the highest estimate of *n* is 679. If a child does not repeat a given sentence *n* times, noisy feedback cannot account for how the child eliminates the error. However, aside from formulas and routines, most of children's speech is not repetitive. Children repeat few sentences, even ten times, let alone 85. For example, Pinker (1989) searched 86,332 child utterances taken from transcripts of spontaneous speech for errors in which a child misused the argument structure of a dative (e.g. *Don't say me that*). None of these errors was repeated anywhere near 85 times: one child (Eve) made eleven errors with the verb *write*, no other error was repeated more than three times, and thus noisy feedback is almost certainly too weak to account for how children eliminate their errors.

Children might follow strategies other than the one I have outlined, but none of these can reduce n without increasing the error rate. For instance, children might be thoroughly productive and assume that every sentence is grammatical unless the sentence receives frequent and repeated correction. Following this strategy, children would overgenerate implausibly often, repeating ungrammatical sentences far more often than they are actually observed to do (Marcus et al., 1992). Children also cannot be completely conservative and assume that all sentences are ungrammatical, since they productively create sentences they have never heard (e.g. Pinker, 1989). In any case, if children utilize noisy feedback before repeating a sentence enough times, their risk of errors must go up, because there is only a limited amount of information in each reply instance, and the only way to get more information is to repeat their sentences again and again in effort to increase the reliability of the information contained in parental replies.

Combining noisy feedback from different sentences I have assumed that children determine grammaticality on a sentence-by-sentence basis. Could children instead determine the grammaticality of entire classes of sentences on the basis of negative evidence? For instance, a child trying to test whether *The boy eat apples* is grammatical might combine feedback provided to that sentence with feedback provided to "equivalent" sentences such as *The girl eat apples*, in order to overcome the inherent weakness of noisy feedback.

Combining feedback provided to some set of equivalent sentences requires a child to determine which sentences are "equivalent"; determining equivalency, however, can only be done properly after the child knows which factors affect a sentence's grammaticality, defeating the very purpose of using noisy feedback. Sometimes two sentences similar in structure, differing by no more than a word, are both equally acceptable in a language; for example, *A boy ate some apples* and *a girl ate some apples* are both grammatical, and *some a boy apples ate* and *some a girl apples ate* are both ungrammatical. Often, however, two apparently quite similar sentences are not equally acceptable in the target language. For example, *a boy ate some apples* is grammatical but *a boy ate much apples* is ungrammatical. A child who combined the feedback provided for these two sentences would err, as would a child who combined feedback for **fill the water in the bowl* with *pour the water in the bowl*, or *John slapped Bill* with **John hitted Bill.* The rules governing which sentences are and are not equivalent, with respect to feedback, are the very rules of the grammar the child is trying to learn. If feedback is to be the means by which children form equivalence classes then children's use of feedback cannot rely on the pre-existence of those equivalence classes. There is no doubt that children do form equivalence classes, but negative evidence does not explain how children form them.[5]

Learning the probabilities of noisy feedback. The model presented above assumes that the child has perfect knowledge of (pr{gr} and pr{ungr}). If the child's knowledge of the reply probabilities is poor, errors of falsely accepting ungrammatical sentences or else errors of falsely rejecting grammatical sentences become more likely. For instance, if pr{gr} .12 and pr{ungr} = .2, but the child incorrectly hypothesizes that pr{gr} = .2 and pr{ungr} = .4, and the child utters a grammatical sentence multiples times and receives p(obs) = .2, the child will conclude falsely that the sentence is grammatical.

Given individual differences between parents (to be discussed below), it is unlikely that the probabilities of feedback are specified innately. How, then, could the child learn

[5]The child learning from positive evidence alone may seem to face the same dilemma -- how to form the equivalence classes without knowing them. The answer proposed by Chomsky and others is that the child has innate knowledge which constrains the inductions by which the child may form equivalence classes.

the reply probabilities? The child cannot simply record how many of her ungrammatical versus grammatical sentences elicit instances of some reply. Children could safely assume that their parents speak grammatically, but by hypothesis they do not know which sentences are ungrammatical. A child might construct minimal pairs by comparing the parent's sentence with a sentence that is likely to be ungrammatical. But this begs the question: how would the child know which sentences are likely to be ungrammatical? In sum, it is unclear how children could learn the probabilities of feedback.

Some researchers seem to want to eliminate language-specific learning mechanisms in favor of parental feedback on the grounds that models that depend on parental feedback might be simpler, but consider what new mechanisms they must propose. The child must recognize the (subtle) feedback reply types and discriminate the useful reply types from the noise, and then make use of them. Furthermore, the child must learn (at least approximations of) the probabilities of feedback, induced in some unspecified manner. Though a child might solve all these problems, it is hardly clear that models of language acquisition that depend on feedback present simpler models than those based on internal, linguistic-specific constraints.

Summary I distinguished three types of negative evidence: complete, partial, and noisy feedback. There is no evidence that either complete feedback, negative evidence provided following all ungrammatical sentences and no grammatical sentences, or partial feedback, negative evidence provided for some ungrammatical sentences but no grammatical sentences, exist. Noisy feedback, the only attested type of feedback, is negative evidence provided for some grammatical sentences as well as some ungrammatical sentences. Under conditions of noisy feedback (unlike complete feedback and partial feedback), a single reply instance cannot guarantee that a given sentence is ungrammatical. Under certain reasonable assumptions about feedback probabilities and acceptable error levels, a child would have to repeat the sentence in question, *verbatim*, at least 85 times before deciding whether it is grammatical. Furthermore, to use noisy feedback, the child must solve the difficult bootstrapping problem of determining the probabilities of reply instances for both grammatical and ungrammatical sentences. Theories of language acquisition that require noisy feedback are thus plausible only to the extent that it is plausible that children notice noisy feedback, have mechanisms to make use of it, and repeat sentences at least 85 times each.

Generality and availability of noisy feedback

Noisy feedback is interesting only if it can eliminate the need for specific internal mechanisms, but these mechanisms are only eliminable if noisy feedback is necessary for language learning. If noisy feedback were merely helpful for language learning, but not necessary, internal mechanisms would still be needed to account for language acquisition. In this section I will develop and apply criteria, based in part on those discussed by Grimshaw and Pinker (1989), Pinker (1989), and Morgan and Travis (1989), for determining whether there is evidence that any type of feedback is a necessary component of language acquisition.

Criteria

Does every child receive noisy feedback? Since every child learns language, a particular type of parental feedback can be a precondition for learning language only if it is provided to every child. Many types of feedback may be *sufficient* for language acquisition, but if any child manages to learn language lacking that particular type of feedback, then that particular type of parental feedback cannot be necessary for language acquisition. (Note that every child receives positive evidence.) Because between-subject differences are crucial, appropriate statistical analyses of patterns of noisy feedback should examine patterns of feedback within individual parent-child dyads. Combining data from several children may create two types of averaging artifacts: First, aggregating data from multiple children may hide differences between the patterns of feedback provided to those children. Suppose one parent provided reliable noisy feedback, but another parent did not provide reliable feedback. Combining the data from these two parent-child dyads, one might falsely conclude that the pattern of feedback is available to both children. To avoid such problems, a more fine-grained analysis is necessary. Second, when studies average raw data[6] from several dyads, patterns of noisy feedback can emerge *that no child actually received.* For example, if 80% of Johnny's utterances are grammatical, and Johnny's mother repeats a random 75% of his utterances, but 60% of Billy's utterances are grammatical and his mother only repeats half of Billy's sentences, at random, then combining the same amount of data for both children, will lead to the (false) conclusion that 64% of grammatical utterances elicited repetitions, compared to only 58% of ungrammatical utterances; see Table 3.

Table 3:
How patterns of feedback could be an artifact of averaging data from several children.

	Grammatical	Ungrammatical
Johnny	60/80=.75	15/20=.75
Billy	30/60=.5	20/40=.5
Average	90/140=.64	35/60=.58

Note:
Number of utterances repeated out of 140.

Do parents provide noisy feedback until errors stop? If certain aspects of language acquisition depend on parental feedback, parents must provide feedback until children have finished acquiring those aspects language. Moreover, if feedback declines or disappears with age, combining data might suggest falsely that noisy feedback is available, even if no child actually receives such feedback. Suppose that a parent consistently repeats every utterance of his or her child at age two but by the time the child is three the parent no longer repeats many of the child's utterances. If at age two, the child produces 80% ungrammatical utterances, but by age three the child only produces 60% ungrammatical

[6] Averaging percentages rather than raw data would prevent the emergence of entirely spurious reply contingencies but could still conceal important individual differences.

sentences, then, as shown in Table 4, one might conclude falsely that the parent repeats a higher proportion of ungrammatical sentences than grammatical sentences, *even though no parent ever replies differentially to grammatical and ungrammatical sentences.*

Table 4:

How patterns of feedback could be artifact of averaging over age.

	Grammatical	Ungrammatical
2 year old	20/20=1	80/80=1
3 year old	0/40=0	0/60=.0
Average	20/60=.33	80/140=.53

Note:
Number of utterances repeated out of 100.

The misleading conclusion results from (a) the parent's repeating fewer of the older child's sentences and (b) the older child's greater proportion of grammatical sentences. Averaging data across different children of different ages may cause similar artifacts.

Do parents provide noisy feedback for all kinds of linguistic errors? Language has several components or levels of representation, including phonological, syntactic, and semantic representations. There is empirical evidence that parents reply to different types of errors in different ways. For example, Brown and Hanlon (1970, p. 48) concluded that the "bases for approval and disapproval ... are almost always semantic or phonological." To show that negative feedback is necessary for the elimination of errors of some kind, phonological, morphological, syntactic, or semantic, one must provide evidence that adequate feedback corresponding to that specific kind of error exists.

Do parents provide noisy feedback for different types of grammatical constructions? Parents might provide feedback for some syntactic constructions, but still fail to provide feedback for other syntactic constructions. A child who received parental feedback for, say, agreement errors (e.g., *He eat the candy*), but not for improper locative verb alternations (e.g., *fill the little sugars up in the bowl*), could not learn which locative alternations are permissible on the basis of parental feedback. Parental feedback must be available for all grammatical constructions for which it is posited to correct.

Summary. To determine whether a particular type of noisy feedback could be necessary for unlearning errors, there are four criteria: (1) The reply type must be available to all children. (2) The reply type must be available throughout acquisition. (3) The reply type must be available for errors in each component of language. (4) The reply type must be available after all types of errors within a given component. Any pattern of noisy feedback that is necessary for language acquisition would meet all of these criteria. Positive feedback meets all of them: all children, at all ages, receive positive evidence for all types of representation and nearly all types of constructions. Complete feedback, if it existed, would also meet these criteria.

Application of Criteria

Does every child receive noisy feedback? Many patterns of feedback were inconsistent across dyads within the studies that reported them, as indicated by asterisks in Table 5. Furthermore, many of the remaining effects failed to replicate in other studies.[7] The rightmost column indicates that only one pattern of feedback was provided to all children studied: Recasts. Bohannon and Stanowicz found that, on average, parents provided more Recasts after ungrammatical than grammatical sentences. However, though they indicate that parents and non-parents replied differently, it is unclear whether every child received more Recasts after ungrammatical sentences: even Recasts may not be consistent across the children they studied.

---

[7]An anonymous reviewer suggested that differences social or economic status might account for some of the variation in feedback.

Table 5:

Availability of particular types of feedback across all children studied.

	Found By	Not Found By	Did all children receive the same type of feedback?
Explicit Approval	.	BH, HPTS, P	No
Non Sequiturs	.	BH	No
Repetitions	HPTS	DPS*	No
Imitations (or Exact Repetitions)	DPS, BS	P, MT*	No
Expansion	P, BS	DPS*, MT	No
Recasts	BS		Yes
Topic Extensions/ Moveons	DTS	P*, MT*	No
Clarification Qs	DPS, BS	MT	No
Confirmation Qs.		P, MT	No

Notes: Morgan and Travis distinguished Partial and Exact Imitations. Exact Imitations were consistent across parents; Partial Imitations were not. Penner found that confirmation questions were more common after ungrammatical sentences; these effects, however, were not statistically significant. * indicates that the effect found was inconsistent across children within that study BH: Brown and Hanlon HPTS: Hirsh-Pasek et al. DPS: Demetras et al. P: Penner BS: Bohannon and Stanowicz MT: Morgan and Travis

The studies that allow examination of individual differences show that different parents provide different, even conflicting patterns of feedback. Demetras et al. found that some children receive certain reply types, such as Expanded Repetitions, more often after ill-formed sentences than well-formed sentences.[8] However, other children receive the same reply types more often after *well*-formed sentences. Penner noted that "although parents repeated correct child utterances slightly more frequently than incorrect child utterances, the pattern was not consistent for all parents." Morgan and Travis, studying three subjects, Adam, Eve, and Sarah, (from Brown, 1973), found that, for each set of parents, the contingencies between parental replies and the grammaticality of children's utterances were different. For example, while Adam's parents asked more Clarification

---

[8]Demetras et al. presented no inferential statistics, and many patterns of feedback are different for different parent-child dyads. Thus contingencies within individual dyads may be due to chance or sampling error.

Questions after his well-formed wh-questions, Eve and Sarah's parents were more likely to ask Clarification Questions after their children's *ill*-formed wh-questions. If the reply types are contingent on grammaticality for some children, but contingent on ungrammaticality for other children, would some children eliminate *goed* in favor of *went*, while other children eliminate *went* in favor of *goed*? Finally, Hirsh-Pasek et al., Penner, and Bohannon and Stanowicz all combined data from many children in their analyses[9], leading to potential statistical artifacts. Data from children who received no feedback, when averaged with data from children who did receive feedback, might incorrectly suggest that feedback is available to all children. Moreover, as shown above in Table 4, it is even possible that some alleged patterns of feedback were not provided to any children. In fact, there may be no type of feedback that is provided universally. Gordon (1990) argued that in the Piedmont Carolinas, and possibly in other cultures, parents rarely speak directly to their children. The existence of cultures in which parents do not provide any feedback would prove that children can learn from positive evidence alone and that no form of feedback can be a necessary precondition to learning language. Bohannon et al. (1990, p. 224) criticize Gordon's example as being "an anecdote ... [that] hardly constitutes believable counterevidence." In fact, Gordon's example is based on a well-documented ethnographic study from Heath (1983), from which I quote: Trackton adults do not see babies or young children as suitable partners for regular conversation.... [U]nless they wish to issue a warning, give a command, provide a recommendation, or engage the child in a teasing exchange, adults rarely address speech specifically to very young children. (p. 86)

Many other cultures with markedly different patterns of parent-child interaction exist. Ochs and Schieffelin (1984) discuss two such cultures. (See also Eisenberg, 1982; Pye, 1986.) In Kaluli, extensive adult modeling is reported, along with a

---

[9]Penner provided information about the reliability of patterns of feedback across parents, but the other two studies do not. Bohannon and Stanowicz only discuss differences between parents, as a group, and non-parents, as a group.

lack of expansions -- apparently without any evident consequences for either rate or success of acquisition:

> In addition to instructing their children by telling them what to say in often extensive interactional sequences, Kaluli mothers pay attention to the form of their children's utterances. Kaluli correct the phonological, morphological, or lexical form of an utterance, or its pragmatic or semantic meaning. (p. 293)
>
> . . . Rather than offering possible interpretations or guessing at the meaning of what a child is saying, caregivers make extensive use of clarification requests such as 'huh?' and 'what?' in an attempt to elicit clearer expression from the child. . . . However, caregivers do not elaborate or expand utterances initiated by the child. (p. 294)

In the case of Samoan, parental feedback is minimal--again, with no apparent consequences for acquisition:

> [C]aregiver speech is largely talk directed *at* the infant and typically caregivers do not engage in 'conversations' *with* infants over several exchanges. . . . When a small child begins to speak, he or she learns to make his or her needs known to the higher ranking caregiver. The child learns not to necessarily expect a direct response. (p. 296)
>
> Procedures for clarification are sensitive to the relative rank of conversational participants in the following manner. . . . If a low status person's speech is unclear, the burden of clarifications tends to be placed more on the speaker. [This] situation applies to most situations in which young children produce ambiguous or unclear utterances. Both adult and child caregivers tend not to try to determine the message content of such utterances by, for example, repeating or expanding such an utterance with a query intonation.... A caregiver may choose to initiate clarification by asking 'What?' or 'Huh?' but it is up to the child to make his or her speech intelligible to the addressee. (p. 298)

Even if Bohannon et al. concede such examples, they still argue that The absence of a particular form of feedback in a particular community does not belie its utility for those children who do receive it, nor does it mean that no form of feedback is necessary for language learning to proceed normally. (1990, p.224)

Although it is true that some form of feedback may be sufficient for those who receive it, any type of feedback that is not available universally cannot be a necessary type of feedback. (This does not eliminate the possibility that some other form might be necessary.)

Could different children use different reply types to learn language? In that case, each child would need to determine which types of feedback his or her parents provide. One might argue that children attend to all possible reply types, and that the reply types that their parents do not use in a manner contingent on grammaticality have no net effect on the child's grammar. But if, for example, some parents use Repetitions to indicate grammaticality whereas other parents use Repetitions to indicate ungrammaticality, then each child would have to determine whether his or her parents' Repetitions correlate positively, negatively, or not at all with grammaticality.

Determining the status of possible reply types raises three problems. First, the child can only calculate a given reply type's status if the child knows which sentences are grammatical. But this is precisely the knowledge the child is trying to acquire. Second, the bootstrapping problem — determining the probabilities for a particular reply instance without knowing which sentences are ungrammatical — is worse if the child does not know which reply types are and are not contingent on grammaticality. As the child tests more reply types for possible sources of feedback, the child's chance of falsely concluding that some reply type is contingent on grammaticality increases (just as running many *t* tests increases the chance of a Type I error). Finally, learning which reply types are useful may require the child to have an unrealistically large memory and unrealistically powerful processing capacities, because the child would have to record whether each possible reply instance occurred for every sentence or at least for enough sentences to discover reliably the reply instance probabilities of a given type of feedback for a grammatical versus an ungrammatical sentence.

In sum, the only plausible way that some pattern of noisy feedback could be necessary for language acquisition is if every parent provides the same reply type to *every* child, and the child is predisposed to expect that reply type to be contingent on grammaticality. But the empirical evidence, both within the Discourse Studies and across cultures, strongly suggests that no single pattern of feedback is available to all children, and hence noisy feedback is unlikely to be necessary for language acquisition.

Do parents provide noisy feedback until errors stop? Although several studies appear to show that noisy feedback is available to two year olds, they also show that noisy feedback diminishes rapidly or even disappears altogether. Hirsh-Pasek et al. found that noisy feedback *disappears* with age: "...sensitivity to well-formedness is only apparent among [parents of] 2-year-olds. The same pattern of results failed to emerge at other ages" (p. 86). Penner studied two groups of children, and the older group (mean age = 3;0) received weaker feedback on all measures than the younger group (mean age = 2;0). Morgan and Travis found that parental feedback provided to Eve declined significantly from the start of data collection (1;10) to the end of data collection (2;3). Neither Adam nor Sarah showed significant parental reply contingencies after age four. In sum, there is no evidence that children over 4 receive any noisy feedback.[10]

Yet children continue to learn language after feedback has disappeared. Many types of linguistic errors persist throughout the preschool years. Sarah said *I maked it with water*

[10]Furthermore, Bohannon and Stanowicz averaged data from different children of different ages and thus some alleged patterns of feedback might be the result of averaging artifacts.

at age 4;5 (Brown, 1973). She makes these inflectional errors at a greater rate at age 5 than at age 2;6 (Marcus et al., 1992). These errors continue into at least first grade and perhaps as late as 9 or 10. Similarly, locative errors, such as Mark's *And fill the little sugars up in the bowl* (from Pinker 1989) occur until at least age 7 (Bowerman 1988). Thus noisy feedback, given its decline over time, is extremely unlikely to account for the unlearning of many of these errors.

Do parents provide noisy feedback for all kinds of linguistic errors? Hirsh-Pasek et al., Demetras et al., and Penner collapsed errors of many linguistic components (e.g., syntax, phonology, and semantics), making it impossible to distinguish whether parents provide feedback for all types of errors or only for some types of errors. Hirsh-Pasek et al. and Demetras et al. both collapsed errors of phonology (e.g., *evelator* instead of *elevator*) with syntactic errors in which words were misplaced or left out entirely. Penner combined morphological errors (e.g., *I ringed the bell*) with syntactic errors (e.g., *I rang bell*), and perhaps phonological errors as well.[11] Bohannon and Stanowicz distinguished phonological errors from syntactic errors but collapsed morphological and syntactic errors. These studies thus shed little light on whether noisy feedback is provided for different types of errors.

The very limited data available suggest that it is unlikely that parents provide adequate feedback for many types of errors. Bohannon and Stanowicz found that parents are much more likely to correct semantic errors than syntactic errors (88.6% versus 35.9%). Morgan and Travis, the only researchers to separate particular types of grammatical errors, found that none of the children they studied received *any* statistically valid reply types contingent on the grammaticality of specific types of linguistic errors, after age 4. In sum, there is little evidence that parents provide feedback for anything other than the truth value

---

[11]It is unclear from Penner's text how phonological errors were treated, but Bohannon and Stanowicz (1988) write that "Penner (1987) lumped pragmatic, semantic, syntactic, and phonological errors into a single category of ill-formed children's speech" (p. 684-685).

and phonology of their children's utterances, consistent with the conclusions of Brown and Hanlon (1970).

Do parents provide noisy feedback for different types of grammatical constructions? Only Morgan and Travis (1989) present data that allows construction-wise comparisons. These data suggest that different types of linguistic errors may receive opposite types of feedback. For example, since overall Adam's parents asked more Clarification Questions following ill-formed utterances (e.g., past tense errors) than well-formed utterances, Adam might infer that Clarification Questions indicate grammatical errors. But Adam's parents asked more Clarification Questions after well-formed *wh*-questions (questions containing *who*, *what*, *where*, *when*, *which*, or *why*). If Adam eliminated those sentences to which his parents replied with Clarification Questions, he could reduce his past tense errors only to the detriment of his use of *wh*-questions, or vice-versa. To use the parental feedback in Clarification Questions, Adam would need to learn different response contingencies for different grammatical forms; it is far from clear how Adam could figure this out. Finally, suppose Adam asked a *wh*-question with a past tense form (e.g., *who runned*?) and his mother replied with a Clarification Question. How would Adam determine whether his mother was correcting his *wh*-question syntax or correcting his past tense formation?

Summary

There is no evidence that *any* type of parental feedback is available widely enough to obviate the need for specific linguistic mechanisms. Only one type of noisy feedback (Recasts, examined in only one study) was available to all the children studied. No other pattern is provided consistently. No study showed patterns of complete, partial, or even noisy feedback available for children over age 4 despite the fact that children continue to make errors after this age. Every study except Morgan and Travis (1989) collapses morphological and syntactic errors; several studies even collapse phonological errors with syntactic errors. Furthermore, because of dubious practices of averaging data from different children and from children of different ages, some patterns of feedback may conceal individual differences showing that certain patterns of feedback were not available to some children and hence not necessary for language acquisition. Worse, some alleged types of noisy feedback may be completely artifactual and never provided to any individual child. The existing evidence for noisy feedback thus fails to withstand the weight of careful empirical scrutiny and provides no evidence that noisy feedback is necessary for language acquisition.

Definitional Artifacts

Many Discourse Studies classified Parental replies in a manner confounded with the grammaticality of children's utterances, hence rendering the correlations between parental replies and children's grammaticality possibly artifactual. The coding criteria for parental replies cause two serious problems. First, many types of parental replies cannot be classified until one looks at the preceding child utterance, and these replies are sometimes coded differently depending on the grammaticality of a child's utterance (Morgan and Travis, 1989). For example, some parental utterances might be coded as Repetitions following a grammatical child utterance, but as Expansions following an ungrammatical child utterance. The child says the identical utterance in (5) and (6), but the utterance is deemed an Expansion or Recast in the former, and as a Repetition in the latter.

(5) Expansion or recast

C: The ball falied down.

P: The ball fell down.

(6) Repetition

C: The ball fell down.

P: The ball fell down.

Second, because parents nearly always speak grammatically (99.44% of their utterances according to Newport et al., 1977), certain parental reply categories, such as Repetitions, are biased to occur more often after grammatical child utterances, while other types of parental coding categories, such as Expansions, are biased to occur more often after ungrammatical child utterances. Thus some of the contingencies reported in the Discourse Studies reflect the consequences of coding categories virtually defined in terms of the grammaticality of the child's utterances, rather than being the result of actual parental sensitivity to the grammaticality of the child's utterances. Consider the following analogy. Suppose that I claim that I can control the color of the sky by my choice of which shirt to wear on a given morning. Every morning, an impartial observer records whether I wear a green shirt or a blue shirt, and records Sky Color, *relative* to my shirt color. After gathering data, we find an interaction between Sky Color and Shirt Color: if I wear a blue shirt then the sky may be the same color, but if I wear a green shirt then the sky is never the same color as my shirt. One might then conclude that the sky responds to or is sensitive to, my choice in shirt color. But this conclusion is false: the sky is not sensitive to whether I wear a blue shirt or a green shirt. The problem lies in the coding criteria for the sky's "replies." We code Sky Color only with respect to Shirt Color: . A blue sky is coded as "Same Color" if I wear blue, but the same blue sky is coded as "Different Color" when I wear green. The sky is not sensitive to my shirt's color; the coding categories are. If we instead simply record the absolute color of the sky, there would be no artifactual contingency and we would not be mislead: The sky is never green no matter what shirt I wear, and the sky is equally likely to blue no matter what shirt I wear.

Exact Repetitions

Repetitions, like Sky Color, are classified relationally. There is no sentence that is a repetition independent of a child's utterance. And just as the "contingency" between Relative Sky Color and Shirt Color is confounded with the fact that the sky is never green, the "contingency" between Exact Repetitions and child grammaticality (i.e., Exact

Repetitions follow more grammatical than ungrammatical sentences) is also confounded. The contingency between Repetitions and well-formed child speech is inevitable given that parents speak grammatically. When children speak ungrammatically, there is virtually no chance that their parents will repeat their utterances verbatim. Conversely, when children speak grammatically, their parents may repeat their utterance exactly. Demetras et al.'s Exact Repetitions, Penner's Repetitions, and Bohannon and Stanowicz's Exact Repetitions are all tainted in this way.[12] Similar problems beset the other parental reply categories that showed some (albeit limited) contingency with grammaticality of children's speech. Recasts, Expansions, and Clarification Questions also can only be coded in relation to children's utterances and hence are inherently confounded with children's grammaticality. For instance, the parental reply *the ball fell down* is coded as a Recast if the child speaks ungrammatically (e.g., *the ball falled down*) but is as Repetition rather than a recast if the child speaks grammatically (e.g., *the ball fell down*). Thus a contingency between Recasts and ungrammatical child speech does not demonstrate that parents are sensitive to their children's grammar. Whenever some aspect of a child's sentence is relevant to the definition of the adult's reply, and that aspect is correlated with grammaticality, any correlation between the grammaticality of the child's utterance and the type of adult reply will be confounded.

Possible Objections

Even if the observed patterns of feedback are the result of spurious correlations, the child could still make use of the parental replies. However, parental replies such as Recasts and Expansions appear not to reflect any parental sensitivity to grammaticality. Parents who are insensitive to children's grammaticality cannot give children *negative* evidence (except through explicit metalinguistic statements of the form *Don't say X*); they can only give

---

[12]Hirsh-Pasek et al.'s category of Repetitions differs from Exact Repetitions because they include parental responses which contain small changes.

children *positive evidence*. .Recasts and Expansions may serve as ideal positive evidence, but they do not tell children what is not in the language. Another possibility is that whenever a parent says something differently from the child, the child should assume he or she has made an error. A child following such a strategy would surely err: most parental utterances are different from their children's utterance even when the child speaks grammatically. For instance, a child might say *I want a cookie* and the mother might naturally reply *No, you've already had three cookies.* Discourse is driven by conversation, not implicit language lessons. Parents may choose a different word or construction to emphasize something different, or even change the topic entirely. Children who changed their grammars every time the parent said something different would radically damage their languages.

Summary

Coding parental replies relative to the child's previous utterance, as has been done often in the Discourse Studies reviewed here, leads to the possibility that the resulting correlations do not reflect true parental sensitivity but instead reflect only constant, non-contingent frequencies of certain kinds of child and parental behavior. The only evidence for noisy feedback comes from reply categories that are definitionally flawed. There are some reply types that are not defined relationally, such as explicit correction, and failures of comprehension (Brown and Hanlon, 1970) . But everybody agrees that these reply types are not contingent on grammaticality.

Discussion

There are three serious problems with the position that parents provide negative evidence to help their children learn language. First, if such feedback does exist, it is too noisy to be useful in practice. Second, it is not available to all children, not available at all the relevant ages, and probably not available for many types of errors. Third, reply categories have often been defined relative to the child's preceding utterance, and thus observed correlations between parental replies and children's utterances may simply be artifacts of the coding scheme. Any one of these problems is enough to significantly undermine the position that parents provide negative evidence to their children is significantly undermined. Positive evidence, in contrast, suffers from none of these problems; nor would complete feedback if it existed. If a parent says something once, the child can assume that it is grammatical; the child need not wait for 85 repetitions. Positive evidence is available to every child regardless of age for all types of representation. Further, every parental utterance is positive evidence regardless of the child utterance. If complete feedback (i.e., negative evidence in the sense used by learnability theorists) existed, it would also meet every test I have applied. If parents provided some reply type following all and only ungrammatical sentences, then a single utterance would be a guaranteed test for ungrammaticality, not the 85 times necessary for noisy feedback. Complete feedback would be available by definition to all children at all ages for all types of errors and could be coded independently of children's utterances. But complete feedback apparently does not exist -- nobody has ever claimed that parents correct all and only grammatical errors. The fact that there is independent evidence for internal mechanisms provides a further argument against negative evidence. Consider errors such as Sarah's *I maked it with water.* As children acquire the English past tense system, they sometimes apply the regular past rule (add *-ed*) to irregular stems (e.g., *go*, *make*, or *sing*), thus producing erroneous past tense forms such as *goed* or *maked.* (See Marcus et al. (1992)) If negative evidence is not available, children must stop producing these forms through some internal (possibly linguistically specific) mechanism. Marcus et al. (1992) argue that children follow a principle of Inflectional Blocking that prevents the application of a regular rule (add *-ed* to form the past tense) whenever a child can retrieve an irregular past tense form for some stem. For example, if a child attempting to mark the past tense of *make* retrieves *made*, the regular rule is blocked. If the child fails to retrieve an irregular past tense form, the rule applies, and the child creates an incorrect form, such as *maked.* As children's retrieval of the correct past tense forms improves through positive evidence, overregularization errors disappear. Internal mechanisms, without the aid of negative evidence, could also eliminate errors such as Mark's *I filled the sugars up into the bowl*, or a similar error, *I filled the water into the bowl.* In these "locative" errors, children use verbs in inappropriate syntactic constructions. The child has used the "content" argument (*the water*) rather than the container "argument" (*the bowl*) as the object of the verb fill.[13]In contrast, it is perfectly acceptable to use *pour* with its "content" arguments as the object of the verb, as in *I poured the water into the glass.* Negative evidence might drive locative errors out of a child's grammar, but without negative evidence children would need to eliminate errors through internal mechanisms. Gropen et al. (1991) argued that locative errors result when a universal linking rule, *object affectedness*, is combined with an improper semantic representation of the verb *fill.* The object affectedness rule states that the direct object of a verb corresponds to an object that is specified as affected in some particular way in the semantic representation of a verb. Gropen et al. argued that in the adult grammar, the semantic representation of the verb *fill* does not specify how the content is affected (a glass can be filled by pouring but also by dipping or bailing as well), hence the content argument cannot appear as the object. In contrast, the semantics of *pour* does specify that its content argument is affected in a specific way, namely it must move downward in a stream, and hence it *may* appear as the direct object of the verb. According to this theory, children always follow the universal linking rule, but if they have an improper semantic representation for a verb such as *fill*, which does specify the manner of motion of the content argument (e.g. being poured), then they would allow *water* to be the object. Gropen et al. predicted that children who thought that *fill*'s semantics specified that its content argument was affected (e.g. by selecting pictures of pouring as exemplifying *fill*)

[13]The precise details of the syntax and semantics and their relationship are outside of the scope of this paper.

would be the same children who made syntactic errors such as saying *fill the water into the sink*. On their account, children could learn the correct semantics using only positive evidence (e.g. hearing *fill* used without pouring as when a cup is dipped in a punch bowl) and crucially, once the semantics is learned from positive evidence, the linking rule fixes the syntax automatically, allowing children to unlearn their errors without negative evidence. If negative evidence really were generally available, it is unlikely that the empirical evidence would support the existence of these mechanisms, because they would be solutions to non-problems. But there is evidence for internal mechanisms. For example, Marcus et al. (1992) showed that children are more likely to overregularize verbs to which they have less exposure. As exposure to correct past tense forms increases, overregularization errors decrease, hence supporting the Inflectional Blocking hypothesis. Gropen et al. (1991) found that semantic errors do indeed correlate with syntactic errors; children that make semantic errors with *fill* are the same ones that use the wrong syntactic frames for *fill*, hence supporting the existence of the object affectedness linking rule. These examples suggest that internal linguistic mechanisms are real and that there is no need for negative evidence. Finally, it is important to reject the notion that nativist explanations of language acquisition depend on the lack of negative evidence. Even if perfect negative evidence were available, innate constraints on the generalizations which children make would be necessary because many plausible errors simply never occur. For instance, children never go through a period where they erroneously form yes-no questions by moving the first *is* to the front of the sentence. Although one can turn *The man is hungry into Is the man hungry?*, children never, by a false analogy, turn *The man who is hungry is ordering dinner* into *Is the man who hungry is ordering dinner?* (E.g. Chomsky, 1965, 1980; Crain and Nakayama, 1987). More generally, at every stage of language acquisition -- inferring the meaning of a new word or morpheme, creating a morphological or syntactic rule, or determining the subcategorization frame of a new verb -- the child can make an infinity of logically possible generalizations, regardless of whether negative evidence exists. The child simply cannot cycle through all logical possibilities and check to see what his parents say about each one.

Children do learn languages; but they appear to do so without requiring negative evidence: Complete feedback and partial feedback do not exist. Noisy feedback, the signal that has been claimed to exist, may largely be an averaging artifact and is at best available in a fraction of the circumstances in which it would be needed. Many of the observed patterns of noisy feedback are likely to be averaging artifacts and definitional artifacts, but even if there were no artifacts, the child would need to say a sentence such as *fill the little sugars up in the bowl* more than 85 times to be sufficiently confident that it was an ungrammatical sentence. There is no existing evidence that noisy feedback can account for the unlearning of grammatical errors, and given the inherent weakness of noisy feedback it is also extremely unlikely that any yet-to-be-discovered interactional patterns would be adequate to account for language acquisition. These considerations suggest that the problem of accounting for children's avoidance and recovery from errors in language acquisition is likely to be explained by the nature of their internal learning mechanisms. The specific mechanisms I have discussed, such as Inflectional Blocking and Linking Rules, are not the only ones that have been proposed and may not be empirically correct. But any explanation of language acquisition that depends on parental feedback seems unlikely to succeed, and quantitative examinations of interaction patterns in parent-child conversation are unlikely to shed much light on this important issue.

References

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

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

Berwick, R. C. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

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

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

Bowerman, M. (1983). How do children avoid constructing an overly general grammar in the absence of feedback about what is not a sentence? *Papers and Reports on Child Language Development*, 22.

Bowerman, M. (1987). Commentary: Mechanisms of language acquisition. In B. MacWhinney, ed. *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Bowerman, M. (1988). The 'no negative evidence' problem: How do children avoid constructing an overly general grammar. In J. Hawkins (ed.) Explaining language universals. Oxford: Basil Blackwell.

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

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

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

Chomsky, N. (1965). *Aspects of a theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1980). *Rules and representations.*° New York: Columbia University Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Netherlands: Foris Publications.

Chomsky, N. and Lasnik, H. (1977). Filters and control. *Linguistic Inquiry* , 8, 425-504.

Crain, S. and Nakayama, M. (1987) Structure dependence in grammar formation. *Language* , 63, 522-543.

Demetras, M. J., Post, K. N., and Snow, C. E. (1986). Feedback to first language learners: The role of repetitions and clarification questions. *Journal of Child Language*, 13. 275-292.

Eisenberg, A. (1982). *Language acquisition in cultural perspective*. Berkeley doctoral dissertation.

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

Gold, E. M. (1967). Language identification in the limit. *Information and Control* , 10. 447-74.

Goodnow, J. J. (1955). Determinants of choice-distribution in two-choice situations. *American Journal of Psychology*, 68, 106-116.

Gordon, P. (1990). Learnability and feedback. *Developmental Psychology* , 26, 217-220.

Grimshaw, J. (1981). Form, function, and the Language Acquisition Device. In C. L. Baker and McCarthy, J. J. (eds.), *The logical problem of language acquisition*. Cambridge: MIT Press.

Grimshaw, J. and Pinker, S. (1989). Positive and negative Evidence in language acquisition. (Commentary on D. Lightfoot's 'The child's trigger experience: "Degree-0" Learnability,') *Behavioral and Brain Sciences*, 12, 341.

Gropen, J., Pinker, S., Hollander, M., and Goldberg, R. (1991). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language*, 18, 115-151.

Hays, W. L. (1973/1988). *Statistics*. Fort Worth, TX: Holt, Rinehart and Winston, Inc.

Heath, S. B. (1983). *Ways with words*. Cambridge University Press.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 267-272.

Hirsh-Pasek, K., Treiman, R., and Schneiderman, M. (1984). Brown and Hanlon revisited: Mothers' sensitivity to ungrammatical forms. *Journal of Child Language*, 11, 81-88.

Kuczaj, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, 16, 589-600.

Lasnik, H. (1981). Learnability, restrictiveness, and the evaluation metric. In C. L. Baker and McCarthy, J. J. (eds.), *The Logical Problem of Language Acquisition*. Cambridge: MIT Press.

Lasnik, H. (1989). On certain substitutes for negative data. In R. Matthews and W. Demopoulos (eds.). *Learnability and Linguistic Theory*, 89-105. Dordrecht, Netherlands: Kluwer Academic Press.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., and Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development.* , Serial no. 228.

Matthews, R. J., and Demopoulous, W. (Eds). (1989). *Learnability and linguistic theory.* Dordrecht, Netherlands: Kluwer Academic Press.

Mazurkewich, I. and White, L. (1984). The acquisition of the dative alternation: Unlearning overgeneralizations. *Cognition*, 16, 261-283.

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

Moerk, E. (1991). Positive evidence for negative evidence. *First Language*, 11, 219-251.

Morgan, J. L., Bonamo, K., and Travis, L. L. (1991). Recasts, corrections and negative evidence: Separate and Unequal. Submitted for publication.

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

Newport, E. L., Gleitman H, and Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow and C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.

Ochs, E., & Schieffelin, B. B. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. A. Shweder & R. A. LeVine (Eds.), *Culture theory: Essays on mind, self and emotion*. Cambridge University Press.

Osherson, D., Stob, M., and Weinstein, S. (1986). *Systems that learn*. Cambridge, MA: MIT Press.

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

Pinker, S. (1979). Formal models of language learning. *Cognition* , 7, 217-83.

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

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

Pye, C. (1986). Quiche Mayan speech to children. *Journal of Child Language*, 13, 85-100.

Roeper, T. (1981) On the deductive model and the acquisition of productive morphology. In C. L. Baker and McCarthy, J. J. (eds.), *The logical problem of language acquisition*. Cambridge, MA: MIT Press.

Slobin, D. I. (1975). On the nature of talk to children. In E. H. Lenneberg and E. Lenneberg (Eds.) *Foundations of language development, Vol. 1*. New York: Academic Press.

Stemberger, J. P. (1989). The acquisition of morphology: analysis of a symbolic model of language acquisition. Unpublished manuscript, Department of Linguistics, University of Minnesota.

Valian, V. (in press). *Parental replies: Linguistic Status and Didactic Role*. Cambridge, MA: MIT Press.

Wexler, K. and Culicover, P. (1980). *Formal Principles of language acquisition*. Cambridge, MA: MIT Press.

Wexler, K. and Hamburger, H. (1973). On the insufficiency of surface data for the learning of transformational languages. In J. Hintikka, J. Moravcsik, and P. Suppes (Eds.) *Approaches to Natural language*. Dordrecht, Netherlands: D. Reidel Publishing Co.

Wexler, K. and Manzini, M. R. (1987). Parameters and learnability in binding theory. In T. Roeper and E. Williams (Eds.), *Parameter setting*, 41-76. Dordrecht, Netherlands: D. Reidel Publishing Co.

Zwicky, Arnold (1970). A double regularity in the acquisition of English verb morphology. *Working Papers in Linguistics* 4, Ohio State University.

Appendix: Definitions of Parental Reply Categories

Hirsh-Pasek, Treiman, and Schneiderman (1984)

Strict Repetitions [were coded if the parent repeated] the child's exact wording with the possible exceptions that (a) ill-formed aspects were rendered in their correct grammatical form as in People lives in Florida => People live in Florida. (b) I and you were appropriately interchanged.

Loose Repetitions. [were coded if] (d) content words were replaced with proforms or vice versa ... (e) modifiers were added or deleted ... (f) a phrase of the child's utterance was repeated without the rest of the utterance ... or (g) the child's utterance was embedded with a longer sentence.

Demetras, Post, and Snow (1986)

Approval/correction: These utterances are explicit responses to the child's preceding utterance that either approve or give negative feedback to the child. Words such as yes, no, that's right, signal explicit feedback.

Repetitions. Four types of repetitions [were] coded: exact, contracted, expanded, and extended. The use of deictic forms (e.g. substitution of you for me, come for go) are accepted as repetitions.

Exact Repetitions [were coded if the parental reply was] Exact repetition of what the child said.

Contracted. [were coded if the parental reply was a] Shortening of the child's utterance in any way.

Expanded Repetitions. [represented a] Correction of child's utterance with appropriate syntax or morphology. Child: daddy house/ Mother: Daddy's house.

Clarification Questions These responses refer directly to the child's preceding utterance. Questions that start a new topic are not included. Also any questions requesting knowledge from the child are not included. The different types of questions are as follows:

Wh- Questions ... must start with a wh-word and clarify, otherwise scored as a Move-On.

Occasional Questions are Questions that have a wh-word embedded in them: You went where?

Repetition Questions.These are repetitions (as described above) that have a rising intonation contour at the end of the sentence.

Move-ons (MO): In these utterances the mother uses the same topic or starts a new topic, but does not 'negotiate' with the child for meaning. We infer that the mother understands what the child said, accepts it, and moves on with the conversation.

Penner (1987)

Topic extensions. .. continued the current topic of the interaction, but did not qualify for the placement in another category, such as repetitions or expansions. The topic was defined by the nonverbal as well as the verbal context of the interaction.

Verbal agreement/approval. included confirmations, agreements, or praise using words like "Yes." "That's right", "Good," and "OK."

Expansions: Parents repeated all of part of the utterance and made additions and other grammatical or semantic changes to words and morphemes in the utterance. Parental responses that qualified as expansions were judged to function to expand upon the previous child's utterance (e.g., C: "Ball fall" A: "The ball fell down").

Repetitions: The parent repeated all or part of the child's utterance without adding to the utterance.

No response: The child's utterance was followed by a pause of at least 2 sec. that did not contain a verbal or nonverbal parental response or another child's utterance...

Confirming questions. ... included expansions and repetitions that were accompanied by question intonation. Therefore, this category contained a subset of the responses that were also defined as either expansions or repetitions.

Bohannon and Stanowicz (1988)

Exact Repetitions consisted of verbatim reproductions of the child's entire preceding utterance.

Contracted Repetitions: consisted of the reproduction of a reduced set of elements from the child's preceding utterance.

Expanded repetitions were coded if the adult reproduced major elements of the child's utterance and added new information.

Recasts [were coded] if the adult preserved the child's meaning but replaced elements of the child's utterance (e.g. C: "That be monkey"; A: "That is a monkey")

Clarification questions were ... questions that related to the children's previous utterance without requesting any new information. {but note that not replicable not o mention that they collapsed children, levels} and counted other adult repetitions with rising terminal intonation as exact, contracted, recasted, and expanded.

Morgan and Travis (1989)

No Response. ...if one or more child utterances immediately followed the utterance containing the error and if no utterance in the subsequent adult conversational turn was explicitly related (via complete or partial imitation) to the error-containing utterance.

Expansions: If any utterance ... expanded the child utterance, e.g. (C: Where other stick? P: Where is the stick?)

Imitations: If the adult exactly repeated the child's utterance (reversals in pronouns and deictic terms notwithstanding).

Clarification Questions: If any adult utterance had the force of requesting the child to repeat all or part of the erroneous utterance.

Confirmation Questions: If the adult reply to the erroneous utterance was a yes-no question pertaining to the linguistic content of the error-containing utterance.

Move-on. If no other category is applicable.

Figure 1: Proportion of child utterances of a given sentence which are followed by a reply instance, given *n* = 10. The vertical line, at p(obs) = .36, represents the upper bound of the 99% confidence interval of pr{gr}. The cross-hatched region of ambiguity represents the range of possible proportions of feedback which fall within the 99% confidence intervals of the distributions of parental responses both to grammatical and ungrammatical sentences. The dotted region to the right represents the area of confidence interval of pr{ungr} that is not within the confidence interval of pr{gr}.

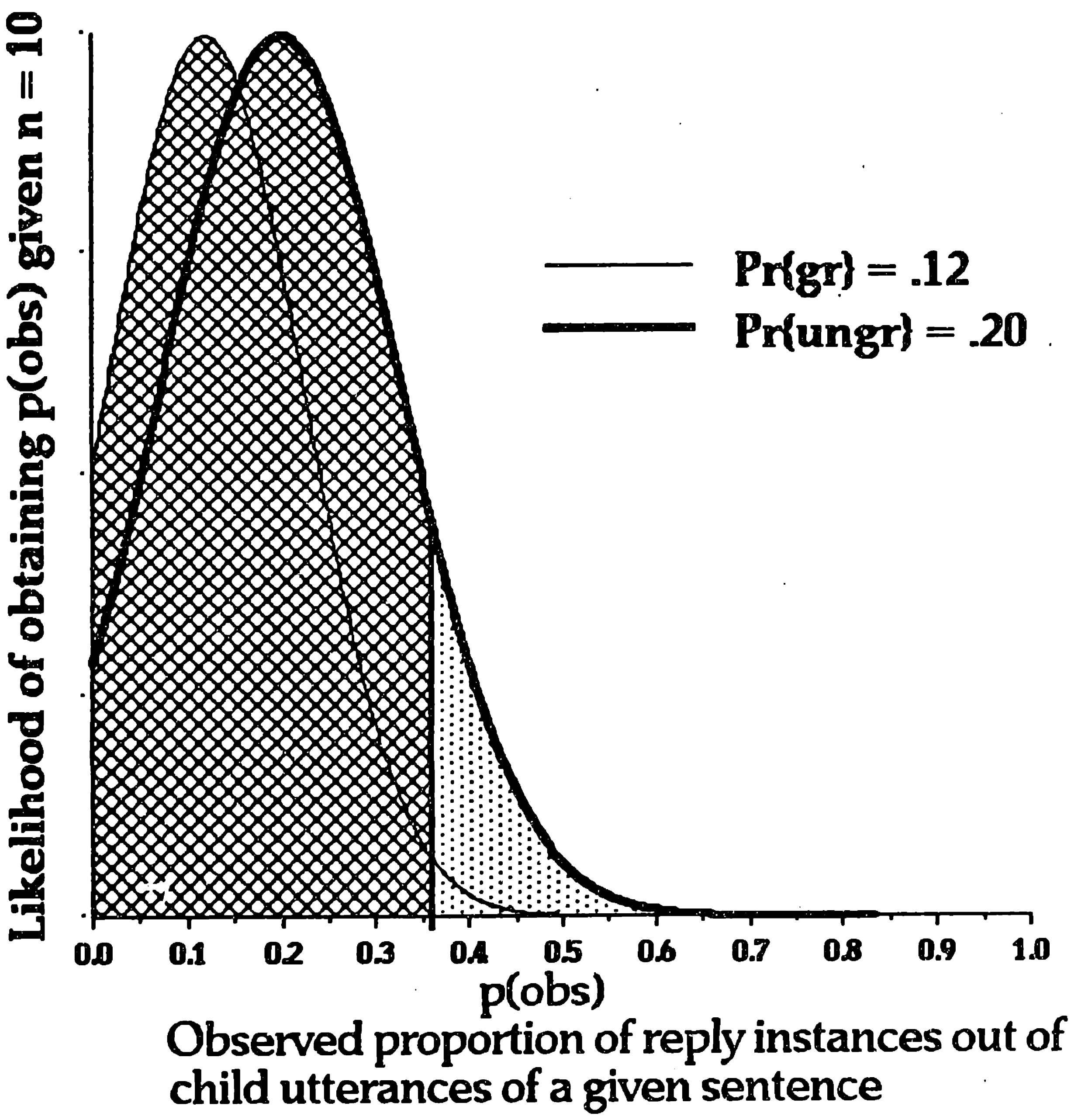
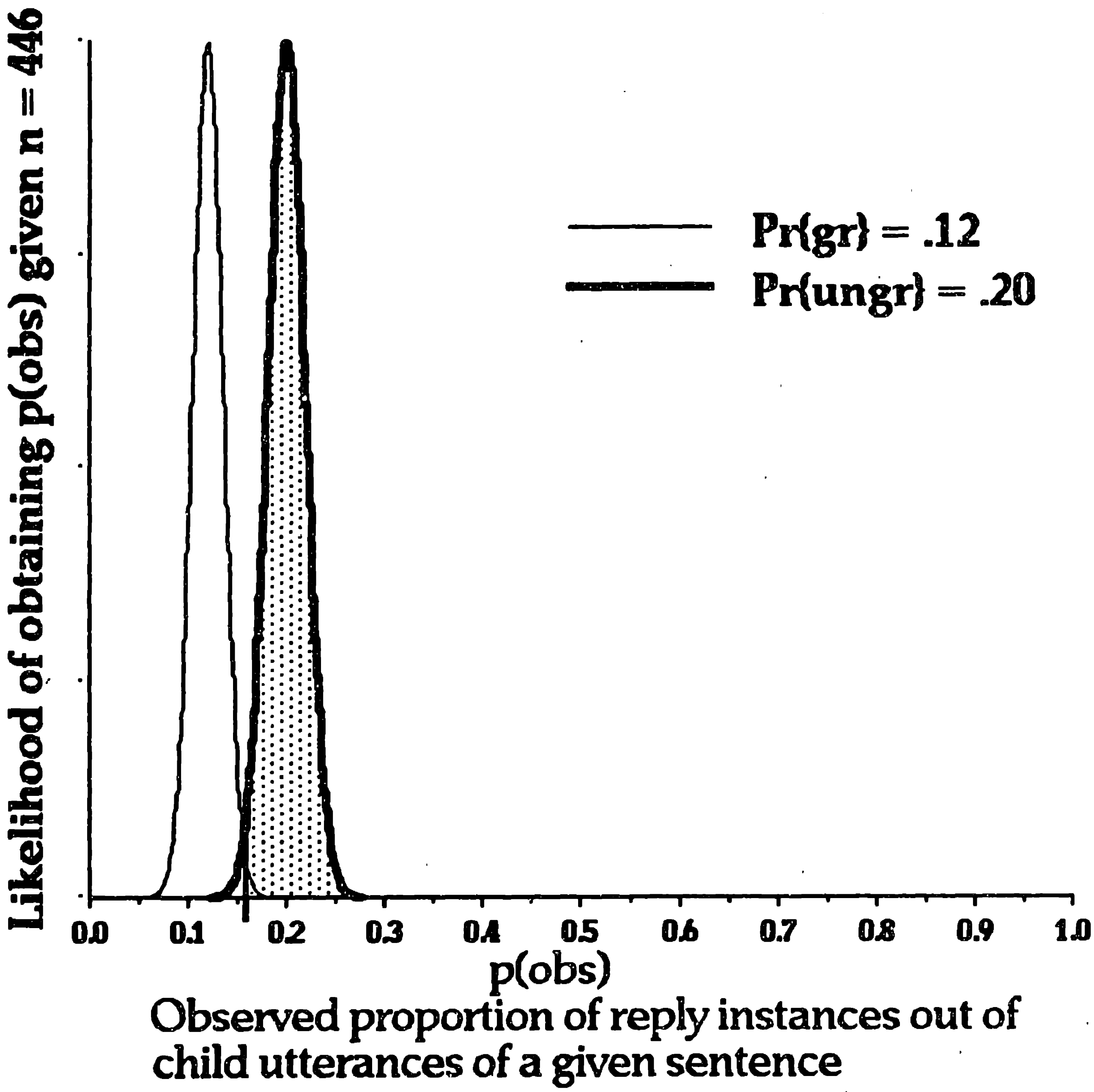


Figure 2: Proportion of child utterances of a given sentence which are followed by a reply instance, given $n = 446$. The vertical line, at p(obs) = .16, represents both the lower bound of the 99% confidence interval of pr{ungr} and the upper bound of the 99% confidence interval of pr{gr}. Thus, the region of ambiguity is zero. The left, dotted, region indicates the confidence interval of pr{gr}. The right, hatched, region represents the confidence interval of pr{ungr}.



PART II

Overregularization in Language Acquisition

7: Introduction

Overregularizations like *comed* and *foots* are among the most conspicuous grammatical errors in child language, and they have been commented upon for as long as language development has been studied (Chamberlain, 1906; Bateman, 1916; Guillaume, 1927; Smith, 1933; Carlton, 1947; Carroll, 1961; Menyuk, 1963; Ervin & Miller, 1963; Ervin, 1964; Miller & Ervin, 1964; Brown & Bellugi, 1964; Cazden, 1968; Slobin, 1971; Brown, 1973; Kuczaj, 1977a, 1981; Slobin, 1978; Bybee & Slobin, 1982a; see Edwards, 1970). These errors are made possible by the fact that English has two ways of creating inflected forms. Most verbs add the suffix *-ed* to their stems to form the past tense, but about 180 exceptional or "irregular" verbs form their past tenses in idiosyncratic ways such as a vowel change (*come-came*), replacement of a final consonant or rhyme (*make-made*, *teach-taught*), substitution of another form (*go-went*), or no change at all (*cut-cut*). Overregularization errors consist of applying the regular pattern to an irregular stem. Since children do not hear these forms from their parents, the errors reveal the operation of a creative process, presumably corresponding to a mental operation implementing the *-ed*-suffixation rule posited by grammarians.

Past tense overregularization is just one kind of error in one peripheral aspect of one component of the grammar of one language. Nonetheless it has assumed a surprising prominence in cognitive science over the past several decades. Overregularization has been offered as the quintessential demonstration of the creative essence of human language (Chomsky, 1959), and of the necessity of explaining cognitive processes by rules and representations rather than by rote and reinforcement (e.g., Brown & Bellugi, 1964; Lenneberg, 1964; McNeill, 1966; Slobin, 1971; Smith, Langston, and Nisbett, in press).

Overregularization has also become famous because of its interesting developmental course, first noted by Ervin & Miller (1963; see also Miller & Ervin, 1964; Cazden, 1968; Pinker & Prince, 1988). The first overregularization errors seem to appear after a period in which children, when they mark tense on irregular verbs at all, do so correctly. Overregularization thus represents a decline in performance in overt tense marking, resulting in a "U"-shaped curve if the proportion of irregular past tense forms that are correct is plotted against age. The nonmonotonicity has been taken as evidence for successive "reorganizations" of the child's linguistic system, reflecting a tendency to ferret out generalizations and to prefer them to exceptional forms which later must be reintegrated in some fashion. As an example of cognitive reorganization triggered by a fondness for regularity, overregularization has ofteneen considered paradigmatic of language development (Slobin, 1973; Bowerman, 1982) and cognitive development (Strauss, 1982; Bever, 1982) in general. It has been used as a metaphor for psychological processes ranging from solving balance-beam problems (Karmiloff-Smith and Inhelder, 1974/1975) to learning computer user interfaces (Grudin & Norman, 1991) to developing expertise in medical decision making (Patel and Groen, 1991).

Recentiy, explaining overregularization has become relevant to the very foundations of cognitive science. Virtually all discussions have implicitly assumed that overregularizations can only be modeled in principle by some explicit representation of a rule in the child's head. Rumelhart and McClelland (1986, 1987) showed this assumption to be false. They devised a computer simulation of an associative network that acquired hundreds of regular and irregular verbs and generalized properly to dozens of new verbs that it had not been trained on. More strikingly, the model appeared to go through a U-shaped developmental sequence, first producing irregular verb forms correctly and later overregularizing them, and seemed to manifest several other effects previously known to characterize children's behavior. But the model had no explicit representation of words, rules, or a distinction between regular and irregular systems; it simply mapped from features standing for the sounds of the verb stem to features standing for the sounds of the past tense form. The Rumelhart-McClelland Model is a prominent representative of the "Parallel Distributed Processing" ("PDP"), "Connectionist," or "Neural Networks" approach, in which cognitive processes are modeled as densely interconnected networks of simple neuronlike units. Its apparent success is commonly seen as a dramatic piece of support for the PDP approach in general, and as posing a severe challenge to rule-based approaches to language and cognition (see, e.g., Rumelhart & McClelland, 1986; McClelland, Rumelhart, & Hinton, 1986; Smolensky, 1988). One reviewer (Sampson, 1987), writing in the *Times Literary Supplement*, called the potential implications of the Rumelhart-McClelland model for the study of language "awesome", because "to continue teaching the subject [linguistics] in the orthodox style would be like keeping alchemy alive."

Pinker and Prince (1988) and Lachter and Bever (1988) have challenged the psychological reality of the Rumelhart-McClelland Model on a variety of grounds. But they praised it for the unprecedented precision of its quantitative predictions about child language data. Indeed, Rumelhart and McClelland were far ahead of the field of developmental psycholinguistics in the quality and quantity of data required for proper tests of their hypotheses. As such they have underscored the absence of systematic quantitative reports of the developmental course of overregularization, its distribution across children, verbs, and time, its relation to the child's vocabulary size, and the lexical factors that cause some verbs to be overregularized more than others. This data gap has left many fundamental questions unanswered.

The phenomenon of overregularization itself is not in doubt, nor is the creative nature of the psychological processes that cause it, which reveals itself in other ways. Children frequently inflect their own invented verbs such as *speeched* (Chamberlain, 1906), *by-ed* (= "went by"; Miller & Ervin, 1964); *eat lunched* (Kuczaj, 1977a), *broomed* (Clark, 1982), and *grained* (Pinker, Lebeaux, & Frost, 1987). In many experiments, beginning with Berko's (1958) classic "wug test," children are given a made-up verb and are then asked to use it in a past tense context, such as "Here is a man who likes to rick. He did the same thing yesterday. Yesterday he ____." The children readily produce appropriate forms such as *ricked*, and when provided with an existing irregular stem, frequently overregularize it (Berko, 1958; Miller & Ervin, 1964; Anisfeld & Tucker, 1967; Bryant & Anisfeld, 1969; Kuczaj, 1978; Derwing & Baker, 1979; Pinker, Lebeaux, & Frost, 1987; Marchman, 1988; Cox, 1989; Kim, Marcus, Hollander, and Pinker, 1991). In other experiments children have been found to judge overregularizations (Kuczaj, 1978) or made-up forms resembling regularly inflected forms (Anisfeld, Barlow, & Frail, 1968; Anisfeld & Gordon, 1968) as acceptable.

Beyond the mere fact that children do more than memorize their parents' verb forms, however, not much is known about the details of overregularization. Nowhere in the literature can one find solid answers to such basic questions as: How often do children overregularize? (1% of the opportunities? 50%? 100%?) At what age do they start? At what age do they stop? What is the developmental curve for irregular verbs, and does it really look like a U? Are all verbs overregularized, or only some, and what factors account for differences among verbs? How is overregularization related to other events in children's language development, such as their vocabulary growth, and the syntax of tense marking? Given how often overregularization has been used to explain other things, it is surprising how little about the phenomenon itself has been documented.

This monograph is an attempt to fill these gaps. Using the large set of transcripts of children's spontaneous speech recently made available by the Child Language Data

Exchange System (ChiLDES; MacWhinney and Snow, 1985, 1990), together with previously published tallies of children's vocabulary and new unpublished data, we document the process of overregularization in quantitative detail. With these data in hand, we will be in a position to evaluate explanations of the psychological processes causing overregularization.

Presumably children are not designed to overregularize per se; their errors are a temporary side effect of a mechanism designed to learn the language. Thus our investigation of overregularization is organized by a research program addressed to the central question of language acquisition research: how children acquire the ability to produce and understand an infinite number of sentences of their native language from the finite sample of speech available in conversations with adults. The "learnability approach" to language development (see Pinker, 1979, 1984; Wexler & Culicover, 1980) seeks to characterize these learning mechanisms; it treats language acquisition as a difficult software engineering problem and attempts to understand how children solve it. The approach dictates the starting point for this monograph (or indeed, for any systematic study of language development). One must begin with an analysis of what is learned, what information is available to learn it, and what kind of computational mechanisms are capable of carrying out the learning task successfully. Such mechanisms define hypotheses about the psychology of the child, which can then be tested and refined against developmental and linguistic data. (As we shall see, overregularization is a prime example of one of the fundamental problems in understanding language learnability: how children avoid or unlearn errors in the absence of parental corrections.) Learnability is not a theory but a research problem, and as such it in no way presupposes the nature of the child's language learning mechanisms (e.g., whether they are symbolic rules or connectionist networks; whether they are specific to language or widely applicable across many cognitive domains). What it does is pick out the central empirical fact to be explained -- that children, at some point, become capable speakers of a language -- and it seeks an explanation in terms of explicit computational operations, not metaphors or impressionistic descriptions. As such, it attempts to characterize the innate mechanisms responsible for human linguistic abilities. This does not single out learnability research as a "nativist" approach, because all explicit theories involving learning must at some point specify the innate mechanisms that do the learning if they are to avoid an infinite regress.

The first substantive chapter of the monograph, then, lays out the logic of linguistic irregularity and overregularization, and proposes a simple psychological hypothesis of how children could learn the irregular system and why they might produce errors while doing so. We outline behavioral predictions relevant to this hypothesis, and the current state of the evidence concerning them.

After a chapter describing the subjects and methods, we present the vital statistics of overregularization. We begin by estimating the overall rate of overregularization, a figure of obvious importance: as we shall see, one's explanation of children's behavior would be very different depending on whether they make overregularization errors almost none of the time or almost all of the time. We then examine whether the rate of overregularization substantially varies across time, children, verbs, and combinations of these sampling units. For example, we test whether an overall steady overregularization curve for a child might actually be a composite of curves for individual verbs each overregularized intensively for a brief period. We also compare the rate of overregularization to the rate of inflecting regular verbs, and define and search for "U-shaped" developmental sequences.

In the next chapter, we examine a hypothesis that underlay the surprising ability of the Rumelhart-McClelland model to display U-shaped development. The hypothesis is that overregularization is triggered by an increase in the proportion of regular verb forms that the child processes during development. We first discuss the complex issue of how exactly one should compare the behavior of network simulations and children. On the basis of this discussion, we correlate children's overregularization rates with changes in the number of regular verbs and the proportion of regular verbs among all verbs in the children's speech, the speech they hear from their parents, and the children's vocabularies. Measuring vocabulary size is a notoriously difficult problem in language development, and we present a novel technique which we also compare to other estimates in the literature.

In the next chapter we examine evidence that points to the development of the entire productive tense marking system, rather than the balance of regular and irregular verbs among the vocabulary items feeding it, as the immediate cause of the onset of overregularization.

The final empirical chapter tests hypotheses about the causes of overregularization by focusing on inherent properties of different verbs that might cause them to be overregularized more or less often. We consider factors related to memory strength (the frequency of an item in parental speech and its similarity to related items in parental speech), the relatedness of a stem to its past form, and the complexity of the mapping from stem to irregular form. This chapter includes a test of a second hypothesis inspired by the Rumelhart-McClelland Model, that regular and irregular marking are computed in a single pattern associator.

In the concluding chapter, we summarize the findings, integrate them within a simple theory, and discuss related theoretical issues.

8:
Theoretical Background

At first glance, the development of overregularization appears to have a straightforward explanation. This traditional explanation, employed both within psycholinguistics and in the other branches of cognitive science that have used overregularization as a metaphor, relies on a dissociation between two psychological processes: rote memory and rule deployment. The rote process memorizes verb forms one by one, and the child can use it at the outset of language development: Children hear their parents say *broke*, so they say *broke*. But the English regular past tense rule is not present at the outset of language development, and the child cannot deploy it until he or she has learned it, presumably by abstracting the regular pattern from a set of regular forms accumulated over time from parental speech and juxtaposed as past and stem forms of the same verb (*walk-walked*, *use-used*, *play-played*, etc.). The young child sticks to correct forms because they are available from rote memory and there is no machinery capable of overregularizing them yet; the older child, possessing the rule, can apply it to irregular stems, resulting in overregularization errors.

The straightforward explanation is inadequate. The problem is that it does not point to any difference between the rule-possessing child and the rule-possessing adult. But there is a difference: children say *comed* and adults don't. If children say *comed* because they possess a regular rule, why don't adults, who also possess the regular rule, say it too? In fact, the standard explanation in terms of a progression between a rote-only child and a rote-plus-rule child does not make any predictions about overregularization appearing at all. It predicts that the younger, ruleless child would fail a *wug*-test (where the child has to provide an inflected form for a novel word like *to wug*), whereas the older, rule-possessing child would pass one. For similar reasons it predicts that the young, ruleless child would inflect familiar irregular verbs correctly (e.g., *broke*) while leaving new ones unmarked (e.g., *stick*), whereas older children would inflect familiar irregular verbs correctly (e.g., *broke*) and overregularize new ones (e.g., *sticked*). But it does *not* predict that children who first use a given verb properly (e.g., *broke*) would later err on that very verb (e.g., *breaked*), which they do (Ervin & Miller, 1963, Cazden, 1968; the phenomenon is documented in more detail in Section 10). Possessing a rule that can be applied to novel verbs, and applying a rule to existing verbs that already have a past tense form and that do not allow the rule, are two different things, and the traditional account does not distinguish them.

For similar reasons, the rule-rote distinction sheds little light on the other arm of the U, in which *breaked* gives way to exclusive use of *broke*. Once children have developed a rule and are overapplying it, how and why do they curtail its use?

Clearly, understanding the developmental course of overregularization requires more than distinguishing between rule and rote: We must examine how the regular rule interacts with irregular items. Specifically, to explain differences between children and adults in how irregular verbs are treated, we must first understand how adults treat irregular verbs. Then we can ask what children might be doing that adults are not. We must also examine the information about irregular verbs available from conversation with adults that the child's language learning process might feed on. With knowledge of what is learned, and what information is available to learn it, we can develop a theory of how the learning works, and why it might result in errors in intermediate stages.

In this chapter we review the current state of knowledge on these issues (deferring discussion of the Rumelhart-McClelland model to Chapter 11), and will develop a hypothesis that appears adequate to account for both the development and the cessation of overregularization. The hypothesis is extremely simple, uniting a standard proposal from formal linguistics -- that in the human language system, irregular memorized items block the application of regular rules -- and a standard proposal from cognitive psychology -- that retrieval of items from memory is probabilistic and sensitive to frequency of exposure.

According to this hypothesis, children's language system, like that of adults, is designed so that retrieval of an irregular form suppresses overregularization, but retrieval is imperfect, and when it fails, the regular rule applies as a default, leading to overregularization errors. The hypothesis will lead us to an equally simple empirical prediction: that at all ages, evidence that the child's system is designed to suppress overregularization should be available in the form of low overall rates of overregularization in comparison to utterances containing the correct irregular form. We will then compare this prediction to competing ones in the literature. This will set the stage for our investigation of the basic quantitative facts of overregularization in the following chapters.

Blocking of Regularization by Irregular Verbs in Adults

Let us begin with a deceptively simple question. Why are overregularizations "errors" at all? This is not intended as a prescriptive question about what one ought to say, but as a descriptive question about what adults do say. Adults who speak the most common dialect of American English do not say *breaked*, and judge it as sounding deviant or childlike (Pinker and Prince, 1988; Ullman and Pinker, 1990, 1991).

The answer is surely not that overregularizations are defective in terms of communicative function; the meaning of *breaked* is perfectly clear. In fact any child who is willing to overregularize has a communicative advantage over adults. No-change verbs in the adult language like *cut* and *set* are ambiguous between past and nonpast: *On Wednesday I cut the grass* could mean last Wednesday, next Wednesday, or every Wednesday; *On Wednesday I cutted the grass* could only mean a preceding one.

Another unsatisfactory answer is that adults don't say *breaked* and *cutted* because they have never heard other adults say them. It is unsatisfactory because it assumes that adults would fail a *wug*-test, and clearly they do not; people don't stick conservatively to the past tense forms they heard their parents use, or that they currently hear other adults use. Adults have no reluctance in creating or accepting past tense forms of verbs they never encountered before, such as *John plipped* or *Yeltsin has finally out-Gorbachev'd Gorbachev* (Kim, Pinker, Prince, and Prasada, 1991; Prasada and Pinker, 1991). Indeed new verbs enter the language frequently, such as *snarf* (retrieve a computer file), *scarf* (devour), *frob* (randomly try out adjustments), and *mung* (render inoperable), and their past tense forms do not require separate introductions. In other words, one cannot explain why adults avoid *breaked* by saying "they've never heard anyone else say *breaked*," because adults have never heard anyone else say *snarfed*, either, but they don't avoid *snarfed*.

The Blocking Principle

The problem with overregularizations is not that they have never been heard before; it's that the irregular counterpart *has* been heard. Clearly there is a psychological mechanism that causes the experience of hearing an irregular form to block the subsequent application of a regular process to that item. Thus several linguists have defined this phenomenon in terms of a *Blocking principle*, (Aronoff, 1976; Kiparsky, 1982; see also Pinker, 1984, who called it the Unique Entry principle): an idiosyncratic form listed in the mental lexicon as corresponding to a particular grammatical modification of a word (past tense, in this case) blocks the application of a general rule that would effect the same grammatical modification. Thus *broke*, listed as an idiosyncratic past tense form of *break*, blocks application of the regular past tense rule, pre-empting *breaked*; *geese*, listed as a plural of *goose*, blocks *gooses*; *better*, listed as a comparative of the adjective *good*, blocks *gooder*. Of course, some verbs do have two commonly-heard past tense forms, as in *dived* and *dove*; for them, the regular version, which is idiosyncratic in the very fact of flouting the Blocking principle, would be recorded from the speech input just as if it were an irregular (Pinker, 1984; Ullman and Pinker, 1990).

Now that we have a mechanism that causes overregularizations to be treated as deviant by adults, we can state the developmental problem more precisely: Children do not apply Blocking in the same circumstances that adults do. The question now is, why not?

One obvious possibility is that children might have to learn the Blocking principle, and overregularizations appear during the stage before they have done so. In the following section, we show that this hypothesis is probably wrong and that the source of children's overregularization errors must be sought elsewhere. Before doing so, however, we add a few remarks about the status of the Blocking Principle.

The Psycholinguistic Status of the Blocking Principle

Blocking is not just a restatement of the fact that overregularizations are deviant-sounding, nor is it a general prohibition against synonyms. Rather, it is a principle specifically governing the relations among the inflected versions of a given stem. Every word in a language implicitly defines a matrix or "paradigm" of its grammatically modified forms (first person singular present, first person plural present, first person singular past, and so on); see Carstairs (1987). Blocking dictates that any cell of the matrix for a given word that is filled with an unpredictable specific entry from the input may not also be filled by applying a general rule for that particular combination of grammatical features. It thus can be interpreted as a psychological mechanism that suppresses the operation of the regular rule in these circumstances.[14]

Blocking thus differs from a pragmatic prohibition against synonymy, such as Clark's (1987, 1990) "Principle of Contrast" stating that "every two forms contrast in meaning." First, Blocking only rules out competing morphological variants, and is agnostic about synonyms in general (e.g., it does not rule out *couch* and *sofa*; see Gathercole, 1987, 1989, for recent arguments that languages tolerate true synonyms, and Clark, 1988, 1990, for counterarguments).

---

[14]The Blocking principle, in turn, might be reducible to Kiparsky's (1982) more general "Elsewhere condition": if the conditions licensing one morphological process constitute a proper subset of the conditions licensing another morphological process, then whenever the conditions of both processes are met by some item, only the process with the more restrictive conditions may apply; the other process thus applies "elsewhere." In this case, an irregular form like *broke* applies in the conditions "past tense of the verb *break*" and the regular rule applies in the more general condition "past tense of a verb," so the latter is pre-empted.

More importantly, there are empirical reasons why Blocking per se, not Contrast, is needed to explain the unregularizability of irregular verbs. Neither violations of Blocking nor conformity to it are predictable from the availability of contrasts in meaning. Members of variant pairs such as *leapt-leaped, dived-dove, sneaked-snuck* do not differ reliably in meaning (or do so with near-imperceptible subtlety); they seem to be recorded by the learner in response to the brute fact of hearing them (later we discuss one reason why they arise to begin with). Conversely, verbs like *put, make, give, take, have, come, go, throw* each have dozens of extremely distinct meanings, especially in combination with particles such as *out, up, off, in, out*, and *away*. Nonetheless without exception speakers are not tempted to use the regularized versions of these verbs to express the contrasting meanings; *taked the punishment stoically, gived away the answer, haved a ball, maked out*, and so on, are always impossible. This shows that different past tense forms do not necessarily evoke different meanings, and different meanings do not systematically call for different past tense forms (see Pinker & Prince, 1988, Kim, Pinker, Prince, and Prasada, 1991, and Ullman & Pinker, 1990, 1991, for discussion).

In addition, Blocking just rules out certain kinds of forms in a specific level of grammatical representation; it does not guarantee that an alternative way of expressing a given notion exists, and thus it differs from a communicative principle like Contrast that is designed to provide a speaker with exactly one way to express each notion. In fact, there are special circumstances where Blocking applies but the irregular is ruled out for other reasons and the speaker is left with an expressibility gap. In particular, Blocking applies at a specific level of the mental representation of word forms, not to entire words; roughly, it applies at the level of inflected stems. An entire word may consist of a stem plus a (noninflectional) prefix. For example, speakers of English intuitively perceive *forgo* as containing the stem *go*. Hence they cannot say *forgoed* because *went*, the listed past tense form of the stem *go*, blocks the application of the regular rule to *go*. But unlike Blocking, speakers' sense of the *familiarity* or naturalness of an irregular form is a phenomenon that takes place at the level of the entire word, not the stem.[15] *Forgo* is a relatively uncommon word and uses in the past tense are even rarer, so *forwent*, while not sounding grammatically impossible like *forgoed*, nonetheless sounds stilted and unnatural (Pinker & Prince, 1988; Ullman and Pinker, 1991). The verb thus has no perfectly usable past tense form: The regular past is ruled out by Blocking (applying at the level of the stem), the irregular past is tainted by unfamiliarity (applying at the level of the entire word). This shows that Blocking is a mechanism operates on specific representations of the speaker's grammatical system.

Blocking as a Means of Recovering from Childhood Errors in the
Absence of Parental Negative Feedback

Logically speaking, given the Blocking principle, children could learn that forms such as *breaked* are not English; they would just have to hear their parents say *broke*. Conversely, given the knowledge that forms like *breaked* are not English, they could learn the Blocking principle. (As we saw, not hearing one's parents say *breaked* is not enough, because parents never say *wugged* either, and it is admissible.) Are there any empirical reasons to choose between these possibilities? In this section we suggest that it is unlikely that children acquire Blocking from evidence about which forms are ungrammatical; it is far more likely that they determine which forms are ungrammatical using Blocking. The problem stems from the concept "the knowledge that forms like *breaked* are not English" or, more generally, "evidence about which forms are ungrammatical."

Negative Evidence and Overregularization

---

[15]By "irregular form" here we mean "completely irregular," including the meaning of the combination of the prefix and stem, as in *forgo*, whose meaning is not predictable from the combination of *for-* and *go*. This differentiates it from compositionally prefixed forms whose stems preserve their meanings, such as *retake* or *overdo*.

A significant problem in explaining language acquisition is that children do not receive "negative evidence": feedback from parents indicating, for any string of words they may utter, whether it is a grammatical sentence. Children are not corrected or misunderstood more often when they speak ungrammatically (Brown & Hanlon, 1970), and although it is occasionally suggested that there is weak statistical information about grammaticality in the differential likelihood of parents' repetitions, expansions, or topic changes, there is considerable doubt as to whether such information exists and whether it is necessary or even useful to the child (see Bowerman, 1987; Gordon, 1990; Pinker, 1989; Grimshaw & Pinker, 1989; the literature and issues are analyzed in detail by Marcus, 1991). A lack of negative evidence means that if the child ever develops a linguistic system that generates a superset of the target language (all the grammatical forms in the target language, plus some ungrammatical forms, not in the target language), the parental input cannot tell the child anything is wrong (Gold, 1967; Pinker, 1979; Osherson, Stob, and Weinstein, 1985). To explain how the adult grammar is attained, then, one must explain either how the children avoid generating supersets, or, if they do, how they expunge their errors.

Overregularization errors in particular pose this problem. A child who is producing *breaked* and *broke* is speaking a superset of adult English in this domain, which only allows *broke*. And overregularization errors in particular do not reliably occasion negative parental feedback. Kuczaj (1977a, p. 599) noted that in his investigation the children (especially his son Abe who was the main subject) were not systematically corrected for overregularization errors; the following typical chunk of dialogue, which we have found in transcripts of conversations between Kuczaj and Abe (MacWhinney & Snow, 1985), illustrates his observation:

Father: Where is that big piece of paper I gave you yesterday?

Abe: Remember? I writed on it.

Father: Oh that's right don't you have any paper down here buddy?

Moreover it seems unlikely that children attend to corrections, requests for clarification, recastings, and so on, when they do occur. For example Zwicky (1970), in describing systematic overregularization of participles by his 4 1/2 year old daughter, reports that "six subsequent months of frequent corrections by her parents had no noticeable effect." The following dialogue, from Cazden (1972), gives the flavor of such attempts:

Child: My teacher holded the baby rabbits and we patted them.

Adult: Did you say your teacher held the baby rabbits?

Child: Yes.

Adult: What did you say she did?

Child: She holded the baby rabbits and we patted them.

Adult: Did you say she held them tightly?

Child: No, she holded them loosely.

More precisely, Morgan and Travis (1989) report a quantitative study on the availability of negative evidence about overregularizations, and its relation to children's recovery from such errors. They tabulated the number of overregularization errors and errors in *wh*-questions, and their correct alternative forms, in the speech of the children known as Adam, Eve, and Sarah (Brown 1973), and cross-classified them in terms of whether the utterances were followed by parental expansions, exact imitations, partial imitations, clarification questions, confirmation questions, attempts to move the conversation on, and no response. No consistent contingency was found between errors and parental responses: for Adam, expansions and clarification questions were more likely to follow his ungrammatical sentences; for Eve it was expansions and partial imitations that occurred more frequently following her ungrammatical sentences; and for Sarah all five response categories of parental response showed the *opposite* pattern, occurring more frequently after well-formed utterances. Unless a child can figure out the kind of parent he or she has (e.g., a grammatical-sentence-expander or an ungrammatical-sentence-expander), such feedback is useless (see Marcus, 1991). Moreover, Morgan and Travis showed that as children get older, even this inconsistent feedback signal disappears. Since the errors continue after the parent has stopped supplying potential feedback, unlearning the errors must depend on some other information source.

How Blocking Obviates the Need for Negative Evidence

Presumably this information source is a constraint endogenous to the child's language system. Blocking is just the right kind of constraint: Children do not have to receive direct information that *breaked* is ungrammatical; they can infer it from hearing *broke*. In sum, because the input information needed to learn Blocking -- that forms like *breaked* are ungrammatical -- is unavailable to children, it is unlikely to have been learned from the input, but may be part of the machinery that does the learning.[16]

If children's language systems incorporated a mechanism implementing Blocking (Pinker, 1984) or its equivalent (e.g., MacWhinney, 1978), we would have a straightforward explanation of how they recover from overregularizations in the absence of negative evidence, converging on the adult state. Each time an irregular past tense form is heard in parental speech, the child can record it in the lexicon and the regular rule is thereafter blocked from applying to it. Thus Blocking is consistent with a developmental

[16]Some such assumption has proven unavoidable to many researchers devising explicit models of language learning, working in a variety of frameworks. For example, Anderson (1983) required a constraint similar to Blocking when applying his symbol-processing model of cognition called ACT to the acquisition of morphology (though the principle was not specific to language acquisition but constrained his production-system architecture across all the tasks it was given). Rumelhart and McClelland (1986), in their very different connectionist model, also implemented machinery that carries out a version of Blocking. As we shall see, in their supervised learning paradigm, a special input pathway was designed that presented a single "correct" past tense form representing the parental input, and an error-correcting learning procedure acted to suppress the tendency of the model to produce an output that deviated from it. Then a separate mechanism, the Whole-String Binding network, which contains explicit representations of correct and overregularized forms, sets them in competition with one another to select the form that best approximated the output of the network.

sequence progressing from overregularization to correct performance (the adult state) without negative evidence. Blocking would prevent children from ever generating a superset of English; they would progress from *breaked* to *broke*, never saying both at the same stage.

Evidence Raising Doubts About Blocking in Children's Language Development

Unfortunately, when we examine the full developmental sequence of inflection in children we immediately run into a problem. Blocking of overregularizations explains monotonic improvement, not U-shaped development. That is, the principle only tells us how children might get out of an overregularization stage. It does not explain how they got into it; indeed, it seems to predict that the child should never get into it in the first place. A child who respects Blocking should never allow the regular rule to apply to an irregular form; the irregular would win the competition from the start, and no U-shaped sequence should be seen. If children really go from correct irregulars to overregularizations back to correct irregulars, the hypothesis that Blocking is inherent to children's language system is cast into doubt.

At first one might try to explain the full sequence as follows. Suppose Blocking is just one manifestation of a more general One-to-One Principle, stating only that one of two competing forms must be eliminated, not just that the form witnessed in the input eliminates the one generated by rule. Such principles can be found in Clark's (1987) Principle of Contrast, Wexler & Culicover's (1980) Uniqueness Principle, and Slobin's (1973) Operating Principle of One-to-One Mapping.[17] If the competition between regularized and irregular forms is two-way, rather than one-way, then two successive replacements, each respecting the principle, would define a U-shaped developmental

[17]Clark's Principle of Contrast acts more like Blocking per se when it is combined with her Principle of Conventionality, which forces the child to use the form that is conventional in the speech community whenever competing alternatives arise.

sequence. In the first arm of the U, the listed form is eliminated by the newly acquired rule (perhaps falling under the general phenomenon of "imperialism" of newly-acquired inflectional rules discussed by Slobin, 1973, and MacWhinney, 1978). In the second, Blocking per se, which requires that the form attested in the input always win out over the one generated by a rule, applies. Crucially, in this hypothesis the child still avoids a superset of English at all stages: *broke* is replaced by *breaked* which is in turn replaced by *broke*; at no time may the two forms co-occur.

But this solution is inadequate. Aside from the fact that the imperialism phenomenon (across-the-board preference for a newly acquired rule) is itself unexplained, the empirical picture of U-shaped development it assumes, in which the child actually *loses* early irregular forms, does not seem to be accurate given the (admittedly scanty) data available at present. All the authors that have actually recorded overregularization errors note that these errors coexist with, rather than replace, the early irregulars (Ervin and Miller, 1963; Cazden, 1968; Kuczaj, 1977a, 1981). This is exactly what all the principles dictating unique forms proscribe. Maratsos (1987) points to such coexistence, which can last for months or years, as a reason to doubt the existence of such principles. Children, it seems, do generate a superset of English (they use *break* and *broke* simultaneously), so if Blocking is the explanation of how children eventually abandon this superset, it leaves it a puzzle why they adopted it to begin with, and why they retain it for so long.[18]

Blocking and Retrieval Failure:

A Simple Explanation of Overregularization Errors

We appear to be at an impasse. The mechanism required to explain adult language

[18]Another possible solution, originally proposed by Kuczaj (1977a, 1981) is that children may fail to realize that a given irregular form corresponds to the past tense version of some stem. Rather, they may treat the irregular past as an independent verb, and Blocking would not apply. Errors would cease when the two verbs were united, presumably when the child noticed that they were semantically identical except for pastness and (in most cases) were phonologically similar as well. We discuss this solution in Section 13.

(Blocking) seems to be systematically flouted by the child during a period in which correct and overregularized forms coexist. But the information necessary to learn Blocking -- negative evidence through parental feedback -- does not exist. As a result, the processes causing the appearance and disappearance of overregularization have been shrouded in mystery, and it is tempting to treat it as a qualitative developmental stage that the child enters and exits, driven by some fundamental reorganization. But positing two qualitatively different kinds of machinery, one for children, one for adults, is not exactly parsimonious, especially since no account has been proposed of how the unknown changing machinery leads to the development of overregularization. Is there a simpler alternative?

One of them is proposed in MacWhinney (1978, p. 6-7) and Pinker (1984, p. 194-195). Say children possess a correct irregular in lexical long term memory and represent it as the past of the corresponding stem, but either the content of the memory entry for the irregular, the link to the stem, or both are not accessible 100% of the time. If an irregular past tense lexical entry is not retrieved, it obviously cannot block regularization. If the child intends to mark tense, and possesses a process for inflecting arbitrary stems, the output will be an overregularization.[19] Similarly, if the content of a stored past tense form is retrieved, but without its "past tense" link or feature,

[19]The principal difference between MacWhinney's (1978) and Pinker's (1984) expositions is that Pinker takes the Blocking principle, as it was explicated and justified by linguists to explain adult knowledge, and simply attributes it to the child, whereas MacWhinney introduced it as a specific new claim about the child's morphological acquisition system. Note that MacWhinney (1987) has since disavowed the claims that regularizations are produced by a rule and that they are suppressed by rote forms across the board. He proposed instead that regularizations are produced by a propertywise generalization process like that of the Rumelhart-McClelland model, and that overregularizations are suppressed by specific, individually learned inhibitions of each overregularized form by its irregular counterpart (1987, p. 285, p. 295). See also MacWhinney and Leinbach (in press), and Section 14.

overregularizations consisting of an affixed past stem would result, like *broked* or *wented*.[20]

The blocking-and-retrieval-failure hypothesis is appealing because it is deducible from the very logic of irregularity, supplemented only by an uncontroversial fact about human memory known since Ebbinghaus. What is the past tense form of the verb *to shend*, meaning "to shame"? If you answered *shended* then you have overregularized; the correct form is *shent* (Bybee and Slobin, 1982b). Of course, this "error" is not surprising. Irregular forms are not predictable (that is what "irregular" means), so the only way you could have produced *shent* was if you had previously heard it and remembered it. But you have heard it zero times, hence can't have remembered it. Now, if in two years you were asked the question again and overregularized it once more it would still not be surprising, because you would have heard it only once. Since memory storage and retrieval are probabilistic, with a higher probability of retrieval for items that have been presented to the learner more often, hearing an irregular a small number of times should be only somewhat better than not hearing at all. Thus low-frequency irregulars are inherently prone to overregularization (Slobin, 1971; MacWhinney, 1978; Pinker, 1984).

Children, by definition, have not lived as long as adults. Among the life experiences that one accumulates through the years is hearing the past tense forms of irregular verbs. Many verbs for a given child will be like *shent* for an adult: never heard; heard but not attended to; heard and attended to, but not enough times to be able to recall it on demand reliably. A child should overregularize these verbs, even with a grammatical system identical to adults'. If children's memory retrieval is noisier than adults', they should make these errors even more often, holding number of exposures constant. But regardless of

---

[20]Presumably the fourth logically possible retrieval pattern can occur as well: the existence of a stored past tense form linked to a stem is registered, but its content cannot be recovered, a state like that studied in memory phenomena such as Tip of the Tongue (Brown & McNeill, 1966), Feeling-of-Knowing (Hart, 1965), and deep dyslexia (Coltheart, Patterson, and Marshall, 1980). In such cases we would not see overregularization, because the activation of the irregular memory entry blocks the rule; nor would we see the correct irregular, because its content is temporarily unavailable. What would surface is the unmarked stem.

whether there are quantitative differences between children and adults, there need be no qualitative differences; blocking of a regular rule by retrieval of an irregular stored in memory, and a memory retrieval rate of less than 100%, are sufficient to account for the phenomenon.

The retrieval failure hypothesis predicts that overregularization is not confined to childhood, and indeed, it is not. Adults occasionally make overregularization errors in their spontaneous speech (Stemberger, 1982; see Section 10). Errors occur even more often in experiments where adults must utter past tense forms under time pressure, and for some kinds of irregulars, the errors are more likely with the lower-frequency verbs (Bybee and Slobin, 1982a). Even unpressured language use shows the forces of overregularization on low-frequency verbs. About 35 irregular verbs admit regular past tense forms as more-or-less natural alternatives in casual American speech, for example, *dreamt-dreamed* and *dove-dived* (see Pinker & Prince, 1988, and Ullman and Pinker, 1991, for a list). Stemberger (1989) and Ullman and Pinker (1990, 1991) have found that these "doublets" have lower average non past stem frequencies than verbs that are exclusively irregular. Moreover, within doublets, the frequency of the irregular past tense form correlates significantly with experimental subjects' ratings of the naturalness of the irregular versus the regularized past tense forms: low frequency forms like *slew*, *slunk*, *trod*, *rent*, and *strove* were likely to be preferred in regularized versions. Finally, the hypothesis is consistent with the fact that irregulars in general tend to be high in frequency in English, and that lower frequency irregular verbs in earlier stages of the language (e.g., *geld-gelt*, *cleave-clove*, *abide-abode*) were likely to become regular over time (Bybee, 1985).[21] Low frequency past tense forms are always in danger of not being uniformly memorized in some

---

[21]Bybee examined 33 surviving verbs from three classes of strong verbs in Old English. Fifteen have come through in Modern English as irregular verbs; 18 have become regular. The surviving irregulars have a mean Francis & Kucera frequency of 515 over all their inflectional forms, 137 in the past tense; the regularized verbs have a mean frequency of 21 over all forms, 5 in the past tense.

generation; if so, the verbs, if they remain in the language at all, will become regular. Verbs that survive as irregulars are thus more likely to be high in frequency.

In sum, the hypothesis of blocking with occasional retrieval failure augments the traditional rule-rote distinction to explain the time course of overregularization errors as follows. Very young children have not yet learned that English obligatorily marks tense and that an *-ed* suffixation rule is available to do so. However they can memorize past forms from their parents and occasionally use them; all past tense forms recorded from children will thus be correct ones. Overregularizations appear when the regular rule is acquired. But the previously acquired irregulars do not go anywhere, nor are they ever incapable of blocking overregularization: they just have to be retrieved to be able to do the blocking, and they are retrieved probabilistically. The cure for overregularization is living longer, hearing the irregulars more often, and consolidating them in memory, improving retrievability. Crucially, this account serves to demystify overregularization, requiring no qualitative difference between children and adults during the overregularization period.

Predictions about the Rate of Overregularization

One crucial datum about overregularization is its actual rate as a proportion of the child's opportunities to make such errors. If overregularization occurred at a rate of 0%, then Blocking alone would explain everything. Obviously the rate is not 0%. But surprisingly, there are few hard data as to what it is, and this allows for an interesting empirical test.

Predictions of the Blocking-and-Retrieval-Failure Hypothesis

The blocking-and-retrieval-failure hypothesis predicts that the child's linguistic system is at all times designed to suppress regularization of verbs remembered to be irregular. This suppression of regularization cannot be perfect because the child's memory is not perfect, but it is as good as the child's memory retrieval process. If we assume that children's memory for words, though imperfect, is quite good (the child is, after all, successfully using thousands of words, and acquiring them at a rate of approximately one per waking hour; Miller, 1977), then overregularization should be the exception, not the rule, representing the occasional breakdown of a system that is built to suppress the error. The overregularization rate therefore, while not being 0%, should be as close to 0% as the child's rate of successful memory retrieval permits. Minimally, an observed overregularization rate that is systematically less than 50% and not attributable to any factor confounded with irregular forms would serve as evidence that the child's language system is biased against overregularization in favor of an irregular form when it is available. Blocking effects exactly that bias, and the lower the rate turns out to be (assuming it is less than 50%), the less need we would have for any explanation other than Blocking and retrieval failure.

Children's overall overregularization rate, of course, is a weighted average of rates for different verbs, which themselves are predicted to range from 100% (for a verb never heard or attended to in the past tense, like *shent* for adults) to 0% (for an overlearned verb). But the best-learned verbs are the ones children hear their parents use most frequently, and the verbs that parents use most frequently are also likely to be the ones that the child uses most frequently and hence should be better represented in samples of children's speech than the rarer, hence more poorly memorized, hence more overregularization-prone verbs. Thus the design of the child's linguistic system to block regularization of irregular verbs should be apparent in pervasively low rates of overregularization in children's spontaneous speech samples.

Competing Predictions

Strictly speaking, there are no alternative theories of the learning process in the literature that can be said to make competing "predictions" about children's overregularization rate. (Rumelhart and McClelland do present an alternative theory, but we will save it for Chapters 11 - 13, because its critical predictions do not concern the rate of overregularization.) The relation between data and hypothesis has generally gone in the other direction: researchers have made rough assumptions about what the overregularization rate is, and have drawn conclusions about the nature of the learning process from them. But because few of the researchers cite actual data on the overregularization rate, and because their idealizations of the data contradict one another, the researchers are in effect making competing predictions about what the rate should turn out to be in a large-scale quantitative study such as the one we are about to report. In this section we treat these idealizations as predictions of the approaches they have been taken to support, and compare them with the predictions of the blocking-and-retrieval-failure hypothesis.

A rate of 100% is by far the most common idealization of the empirical picture in the literature. Recall that it underlies the One-to-One Mapping hypothesis, which explains the two phases of overregularization in terms of subsequent replacement of irregulars by overregularizations and overregularizations by irregulars. It also inspired the picture of the child as the exception-hating rule-monger that was imported into other branches of developmental psychology. As such the assumption of a 100% overregularization rate can be found everywhere from textbooks to technical articles to the popular press. Here are examples of each:

> Interestingly, when the general rule is learned, children will often stop using the previously learned irregular form and instead produce a regularized version. Thus a child who had been using the correct past tense of *sing* -- namely, *sang* -- may start using the form *singed* ... instead. Before the correct irregular form is again learned, some children have been known to produce forms that combine the irregular and the regular form. In the case of *sing*, this would be *sanged*. Thus children may proceed through as many as five steps in the acquisition of some

inflections ... No inflection, Adult Form, Overregularization, Transition, Adult Form. (from the textbook *Language Development*, Reich, 1986, p. 148).

At this time [overregularization] the irregular forms that the child had used earlier fade out in favor of the overregularized forms. When the irregular forms later reassert themselves, they have a new status: they are no longer isolates operating independently from their uninflected counterparts and from regular inflected forms; rather, they are integrated into a system, as exceptions to it. (from the theoretical chapter "Reorganizational processes in lexical and syntactic development," Bowerman, 1982b, p. 321).

The errors [children] do make are actually logical overgeneralizations of rules. Instead of "He went," for instance, they may say "He goed" -- to them, a perfectly reasonable past tense of "go." They reject irregularities. When they have learned phrases by rote like "It broke" and "two mice," they will toss them out once they become aware of past tenses and plurals, notes University of California psychologist Dan Slobin. Suddenly, they start using "It breaked" and "two mouses." Says Jill de Villiers, only half joking: "Leave children alone and they'd tidy up the English language." (from *Newsweek* magazine, Gelman, 1986, p. 85).

We find a subtly different idealization in the following, superficially similar statement of the facts, from Eve Clark's article arguing for the Principle of Contrast:

Children are pattern-makers. And when they begin to acquire the inflections that mark tense, for instance, they typically take irregular verbs such as *break*, *bring*, and *go*, and **treat them as if they belonged to the regular paradigm** of *walk*, *open*, and *jump*. So the past tense of *break* is produced as *breaked*, *bring* as *bringed*, and *go* as *goed*... (Clark, 1987, p. 19, emphasis added).

Note that Clark is not claiming that irregular past forms are tossed out or even that they fade; she only says that irregulars are treated as regular verbs. Crucially, most children, while they are overregularizing, do not successfully inflect *regular* verbs with *-ed* 100% of the time. Brown (1973) reports that none of the three children he studied were producing regulars more than 90% of the time for six consecutive hours of samples during the time they began to overregularize, though they achieved this criterion several months later. We examine these data in detail in Section 12; for now assume that children inflect regular verbs during the overregularization years at a success rate of 75%. In that case Clark's statement would be consistent with children overregularizing 75% of the time (or whatever the actual rate of inflecting regular verbs turns out to be). Indeed it is broadly consistent with children continuing to produce correct irregulars some proportion of the 25% of uses that are not overregularizations. That, in turn, would be literally be consistent with the finding discussed earlier that children use both overregularized and correct versions of a verb during the overregularization stage, as long as the overregularization rate was always capped by the successful suffixation rate for regular verbs.

However, many writers have interpreted the finding of coexistence as showing a more extreme deviation from complete overregularization. For example, in reference to an aspect of the development of verb argument structure that did not consist of complete replacement of a correct form by an error, Bowerman (1982, p. 342) amends her earlier summary of overregularization as follows:

> However, it has become clearer in recent years that overregularization is not the all-or-none phenomenon it was once taken to be: Irregular forms rarely drop out, but rather continue to compete with their overregularized counterparts throughout the period of error marking... The relative strength of the irregular and overregularized forms in this competition reflects a complex interplay of factors, such as how long the irregular forms have been part of the child's repertoire

> before their role in a broader system is perceived, how frequently they have been said or heard, whether the "irregular" forms are truly mavericks or belong to minor patterns of their own, and whether the child routinely activates a newly grasped systematicity in the course of sentence construction or perceives it only more passively.

No clear value for the assumed overregularization rate is given, but "relative strength of irregular and overregularized forms" implies in context that for some verbs overregularizations predominate, for others the correct form does. There is no reason to expect that overregularizations or correct irregulars are *systematically* preferred, and given the large number of different biasing factors at work over a large number of verbs and children, we might expect on statistical grounds that they would roughly even out and so we can read Bowerman as basing her theorizing here on the assumption that the overall overregularization rate does not deviate from 50% by an amount large enough to attribute to some pervasive single cause.

Maratsos (1983, p. 763), in his review chapter on language development in *Carmichael's Manual of Child Psychology*, is more explicit about his assumptions about children's overregularization rate:

> Kuczaj (1977b) examined longitudinal samples for one subject, and 6-hour cross-sectional samples for 13 children of different levels of competence. Again, no evidence indicates that *-ed* overregularizations drove out irregular past forms. Instead, even within individual verbs, overregularized and irregular forms alternated, often for periods of months to years. Overregularization, in general, ranged in frequency from .20 to .60 of the children's uses.

It is not clear from Maratsos's chapter what the figures of ".20 to .60 of the children's uses," attributed to an unpublished talk by Kuczaj, refer to, as we shall see in Chapter 10. For now, it suffices to note that Maratsos is suggesting that the typical rate of overregularization is not far from 50% (the literal implication of "alternation"), perhaps 40% (the midpoint of the range he cites). It is also clear that these are the data he considers as a refutation of Blocking. In his review of Pinker (1984), he writes:

> Yet empirically, children do not act as though they have such a solution [blocking], but may alternate between the overregularized *-ed* form and the irregular form for a period of months to years, using both *broke* and *breaked*. ... It is clear that their analysis and resolution of such alternatives is a long-drawn-out tabulational process, not one which quickly seizes upon one or two properties of the language as heard. (Maratsos, 1987, p. 19).

Thus the literature contains a range of idealizations of children's typical overregularization rate, and theories based on them, ranging from 100%, to the rate of regular affixation (say, 75% - 95%), to some aggregate rate reflecting system-wide indifference (perhaps not too far from 50%, or some range of values that includes it such as 20-60%). A rate of exactly 0% is of course never entertained; nor is a rate near 0%, or even greater than 0% but pervasively less than 50%, representing a systematic preference for irregulars. Since we have derived this very prediction from the blocking-and-retrieval-failure hypothesis, there are now clear competing predictions. Thus the first empirical tests in this monograph will involve quantitative analyses of large samples of children's speech in an attempt to estimate children's characteristic overregularization rate. The estimate will be compared against the various predictions and assumptions in the literature and thus will test the corresponding ideas about the psychology of overregularization.

Summary. The mere fact that children develop a regular rule cannot explain why they overapply it to irregular verbs. Adults block the application of regular rules to idiosyncratic memorized items; we suggest that children do, too, but their retrieval of idiosyncratic items from memory, especially low frequency ones, is probabilistic, and overregularizations occur when it fails. The hypothesis explains how children unlearn their errors in the absence of negative parental feedback, and does not make the dubious prediction that overregularizations replace correct irregulars in children's speech. It predicts that overregularizations should generally be rare in children's speech relative to correct irregular past tense forms, contrary to most current views about language development.

9:
Method

Subjects

To assess overall overregularization rates we wanted to examine as large and diverse a set of children as possible. Using the 1990 version of the ChiLDES database and documentation (MacWhinney, 1990), we selected all the unimpaired English-speaking children that met the following criteria: (1) the children were known to speak Standard American English; (2) the transcripts were in CHAT format (allowing efficient and accurate computer searches); (3) the investigator did not include warnings that the transcripts were in a preliminary state and possibly dangerous to use; (4) information was available about the subjects and the circumstances in which their speech was recorded; (5) the transcripts contained at least 10 irregular past tense forms per child (without regard to whether they were correct or overregularizations). Applying these criteria yielded a sample of 83 children, who produced 11,521 utterances containing past tense forms of irregular verbs. Table 1 shows the children, their ages, and their recording schedules. Among these 83 children we focus on the 10 represented by longitudinal samples and on the 15 children with single samples from Hall, Nagy, and Linn (1984). The remaining 58 came from multi-child databases, and we used their data as a replication of the basic findings on overregularization rate obtained from the main sample.

Sarah was a child from a working class background whose parents had high school degrees and whose father worked as a clerk; all the other individual children were from professional families. The children in the group databases are all from middle class, though not necessarily professional, backgrounds. Ten of the children were black (Adam and nine of the children in the Hall, Nagy, and Linn sample). There were 5 boys and 5 girls among the children with individual databases, 10 and 5 from Hall, et al., 8 and 6 from Gathercole; 14 and 10 from Gleason; and 10 and 10 from Warren-Leubecker.

Table 1 about here

To answer questions about longitudinal development and vocabulary size, we need to focus on samples that began before the onset of overregularization and continued long enough for performance in tense marking to approach adult levels. Brown's (1973) Adam, Eve, and Sarah meet this criterion; overregularizations are absent from their early transcripts and their later transcripts extend to Brown's "Stage V" in which most inflections are supplied in their correct forms more than 90% of the time.

There are also extensive longitudinal transcripts for Kuczaj's son Abe (Kuczaj, 1976, 1977a, 1978), but they begin later than those of Adam, Eve, and Sarah, and Abe was already overregularizing in the first, so questions about his onset of overregularization must remain unanswered. However, we examine other aspects of his development, and take advantage of the prodigious numbers of overregularizations and correct irregulars in his transcripts when focusing on individual verbs and their developmental courses.

Finally, when examining the effects of lexical factors, we analyzed overregularization rates for individual verbs from all 19 children and correlated them with various properties of the verbs. Such analyses involve a tradeoff: Individual children often supply too few errors to provide the wide of range of overregularization rates and the wide range of predictor variable values needed for correlational analysis, but aggregate data are in danger of displaying averaging artifacts. Therefore in the lexical analyses we seek converging results from 19 individual children who overregularized, and, where possible, an aggregate measure that combines their overregularization rates.

Procedure

We tabulated all past tense forms of irregular verbs, correct and overregularized, for all of the children. (See the Appendix of Pinker and Prince, 1988, for an exhaustive list of irregular verbs in present-day American English, sorted into subclasses.) Irregular verbs admitting of a common regular alternative in adult speech such as *dream* and *dive* were excluded. Overregularizations included stem+*ed* forms like *eated* and past+*ed* forms like *ated*. We did not search for other possible kinds of error involving the regular suffix such as *sweepened* or *brecked* (for *broke*); such errors, in any case, are quite rare. The actual tabulation procedure differed slightly for the different children.

For Abe, the data were gathered and tabulated by Kuczaj (1976), the boy's father. For each month from age 2;5 - 5;0, Kuczaj recorded the number of times Abe used each of 70 irregular verbs, the number of times he produced the present stem of the verb with *-ed* appended, such as *goed* and *breaked* ("stem+*ed*" errors), the number of times he produced doubly marked pasts in which *-ed* was added to the irregular past form, such as *wented* and *broked* ("past+*ed*" errors), and the number of times he used the stem form in obligatory past tense contexts, such as *Yesterday we go out*.) Kuczaj's tables also contain one occurrence of *beed*, but Kuczaj did not tally other uses of *be*, *was*, or *were*, presumably because they could also be forms of the auxiliary *be*; this would also account for why *have-has-had* and *do-did* are absent. Our analyses of Abe thus omit all three verbs. Kuczaj reported these figures in two ways: in a table listing the number of correct and incorrect forms, summed over verbs, for each month (his Table 18), and in an appendix listing how many times each individual verb, in each of its possible forms, was used each month (his Appendix G). Occasionally there were discrepancies between these two tables that we could not resolve from the text of the thesis, so we relied on tallies of the raw data from Kuczaj's investigation, provided to us by Michael Maratsos. Data from the months 2;5 to 2;7, and for 4 no-change verbs, were absent from these tallies, and for them we used the thesis tables exclusively.

For Adam, Eve, and Sarah, verb usages were tallied on a DEC Microvax II running UNIX. Individual transcript files were combined into single master file for each child. The *freq* program in the CLAN software package (MacWhinney and Snow, 1990; MacWhinney, 1990) counts the number of times every word is used in a particular transcript session, for a particular speaker, and it was run on the individual transcript files and the combined files. The combined frequency list for each child was then edited to include only words that the child may have used as a verb, including words that occur only infrequently as a verb (e.g., *fish, color, ground, milk*). For all such items (i.e., all words that are not exclusively verbs in the child's vocabulary), the Unix utility *fgrep*, which finds matches of alphanumeric string patterns, extracted all the transcript lines they occurred in. Each of the resulting lines was checked by hand, and excluded if the matched word turned out not to be used as a verb. If a word appeared in a single-word utterance, it was excluded; thus, "sleep" or "put" appearing alone were not counted as verbs, but "Adam sleep" or "put Mommy" were included. Verbs repeated in successive sentences such as "I falled down. I falled down, Momma" were counted separately, since children are capable of saying both a correct and overregularized version of a single verb in successive utterances, as in Abe's "Daddy comed and said 'hey, what are you doing laying down?' And then a doctor came...".

For all instances of no-change verbs such as *cut* and *put*, past tense usages were distinguished from present and infinitival usages by hand; when the transcript did not provide information regarding the verb's tense, it was inferred from the context. Contractions such as *gimme*, *gonna*, *I'm*, *it's*, or *doesn't*, were excluded, as were participles such as *broken* or *gone*, and the quasi-modal *used to*. A very small number of mimicked utterances at early ages, regular participles, and irregular participles that are identical to past tense forms may have been included. Intentionally included were verbs that were not very clearly uttered, but were clear enough for the transcriber to have made a reasonable guess, and some slight phonetic variations such as *-in* for *-ing*, particularly for Sarah whose samples were transcribed more narrowly than the others. Brown (1973) notes, though, that all three children's speech was carefully transcribed with regard to the presence or absence of phonetic material corresponding to inflections.

For Adam, Eve, and Sarah, *have, be*, and *do* were included only when they were used as main verbs (i.e., possessional *have*, *have to* unless transcribed as *hafta*, copula *be*, and pro-verb *do*), never when they were used as auxiliaries (i.e., perfect auxiliary *have*, progressive and passive auxiliary *be*, *do* used to form questions, negations, or emphatics). This is consistent with the criteria used by Slobin (1971), Bybee and Slobin (1982a), and Brown (1973; Brown, Cazden, and de Villiers, 1971).

For the other children, and for the Gleason, Warren-Leubecker, and Gathercole collections, we used a Sun Microsystems Sparcstation 4 running under UNIX to tabulate all irregular past tense utterances. Using the 'freq' program we extracted the number of occurrences for each uniquely irregular verb listed in the Pinker and Prince (1988) Appendix, together with all forms ending in *-ed*, with the exception of the no-change verbs, *read* (which is orthographically a no-change verb), and all forms of *do*, *be*, and *have* (i.e., neither auxiliary nor main verb usages were counted). We then isolated all the overregularization errors from this list by removing regular verbs and other part of speech categories. Because we did not check this large collection by hand, we were unable to exclude overregularized participles such as in *the window was broked*. Since we did exclude correct participle forms if they were distinct from the past tense forms (about 60 irregular verbs have this property; see Pinker & Prince, 1988), this can result in an overestimate of overregularization rates, though it would be small. The word *seed* presented particular problems since most of its uses are as a noun rather than as an overregularization of *see*; these were eliminated by hand. Repetitions were counted separately, as with the

Brown children.[22]

The verb *get* is complicated. When adults speaking the standard dialect use *got* in a stative possessional sense, as in *I've got an ice cream cone*), it is the perfect participle of *get* (meaning "obtain"), accompanied by some form of the auxiliary *have*. The meaning is possessional because of the semantics of perfect aspect in English: if the state resulting from obtaining something in the past currently holds, you possess it now (Bybee, 1985). However, if children do not attend to the auxiliary, it would be natural for them to reconstrue *got* as a present tense form meaning "possess," and there are numerous forms like *Look, I got an ice cream cone* that suggest they do often use *got* as a present tense verb. For these usages we would erroneously credit the child with the correct past form of *get*. Kuczaj (1976) noted this problem and used the context to distinguish present from past usages of *got* in Abe's speech. For Adam, Eve, and Sarah, we excluded all forms that were clearly present tense statives. For the other children, our observed overregularization rates for *get* are probably underestimates, and since *get* is a frequent verb, the overall overregularization rates across verb tokens will be, too. As we shall see when the verb is separated out, however, the degree of underestimation is small.

To verify the accuracy of the machine-generated tabulations for the children other than Adam, Eve, Sarah, and Abe, we compared an exhaustive hand tabulation for these four children with machine-generated totals like those used for the other children. The hand-generated totals were calculated by (1) extracting all utterances containing irregulars and all utterances with forms containing *-ed*, not including *have*, *be*, *do*, and the no-changers, and

---

[22]Because the ChiLDES transcripts contain typographical errors and inconsistencies, a handful of past tense forms may have been gone undetected by the automatic search procedure. For example, Michael Maratsos has called our attention to an overregularization rendered as *fell'd* in April's transcripts. We ascertained that such isolated misses are rare by checking for all words spelled with *'d* in the transcripts of Adam, Eve, Sarah, and Abe; none were overregularizations. To be consistent, we did not add April's *fell'd* to her overregularization total, because the transcripts from various children contain a small but unknown number of mistranscribed correct irregulars that our searches also missed, such as *gotalright* and *stucki'm*.

(2) checking these utterances to remove all participles, nouns, and all other nonpast forms. The mean discrepancy between the two estimates of the overall overregularization rates for the four children was less than +0.2 percentage points (maximum 1.2 percentage points), with the machine giving a higher estimate for two of the children and the human giving a higher estimate for the other two. We also compared the overregularization rates derived by the two methods across verb types. The two sets of estimates correlated fairly well: r = .98, .77, .91, and .82 for the 4 children, respectively, and the overregularization rates averaged over types were higher when estimated by machine for two of the children, lower for the other two (never by more than 2 percentage points). Most of the discrepancy can be attributed to small samples for some verbs, where disagreement over a single sentence can greatly affect the overregularization rate for that verb (e.g., if a verb was used twice, once correctly and once incorrectly, its estimated overregularization rate could change from 0% to 100% if one of the sentences is omitted or misclassified). For the children other than Adam, excluding verbs used fewer than 10 times brings the correlations between verbs' overregularization rates calculated by hand and by machine to .98 or greater. Thus in several analyses in this monograph, when it is important to exclude verbs that were produced too few times to yield reliable estimates of their overregularization rates, we use a criterion of a minimum of 10 tokens per verb per child.

For several analyses it is necessary to tabulate the frequency of use of past tense forms in the speech of the adults that habitually talked to the children. In all cases the same criteria for counting verbs were used for adults and children. Although we did not distinguish among the different adults talking to a given child (e.g., the mother, father, or psycholinguists, who came to be treated like family members), we did keep separate tallies for the sets of adults talking to each child, so each child has his or her own adult speech data.

Calculating Overregularization Rates

We defined the child's "overregularization rate" as the proportion of tokens of irregular past tense forms that are overregularizations:

$$\frac{\text{\#overregularization tokens}}{\text{\#overregularization tokens + \#correct irregular past tokens}}$$

Overregularization tokens included both stem+*ed* and past+*ed* forms. Virtually all of children's past tense forms are in past tense contexts (Brown, 1973; Kuczaj, 1976), so there is no need to take into account the semantic correctness of the tense marking. Overregularization rates were calculated over tokens for a given verb for a given child, and over tokens of all verbs for a given child.

Rationale for Calculation of Overregularization Rates

Note that this measure of the overregularization rate excludes no-marking errors (stems in obligatory past tense contexts). This is because no-marking errors could be caused by a failure to try to mark the past tense at all (in which case the stem would be used because in English it is the most frequent and grammatically nonspecific form), and the issues at hand pertain to *how* tense is marked, on those occasions when the child decides to mark it at all, not on *whether* the child decides to mark tense, a logically independent question. For example if overreguiarization were measured as a proportion of all forms in past tense contexts (by adding no-marking errors to the denominator), it would appear that young children's overregularization rates were low, but that could just be because they were usually not trying to mark tense at all, not that they were successfully suppressing errors. And if overregularizations were lumped together with no-marking errors to yield an overall error rate (by adding no-marking errors to the numerator) then overregularization errors themselves would be hidden from view; a given error rate could correspond to nothing but overregularization errors, nothing but bare stem errors, or anything in between. In particular, tests of a regression in development (the left arm of the U-shaped developmental sequence) would be impossible. Cazden's (p. 437) describes the sequence as "no use, followed by infrequent but invariably correct use, followed only later by evidence of productivity"; the distinction between "use" and "correct use" presupposes that correct forms are being considered as a proportion of total marked forms (irregulars plus overregularizations). Thus Rumelhart & McClelland (1986), using the same rationale that we do, plot a quantity corresponding to our overregularization rate in their demonstration of how their model mimics U-shaped development, as shown in the graph which we reproduce as Figure 17. That is, Rumelhart and McClelland assume that the past/non past tense distinction has been mastered independently of the computations performed by their model in deriving past tense forms, and so they simply feed it correct stem-past pairs from the start.

Of course, the decision to count only overtly tense-marked forms because they are clear examples where the child has decided to mark tense does not imply that nonmarked stem forms exclusively represent the absence of a decision to mark tense. Rather, no-marking errors are highly ambiguous, and can arise in at least four ways. First, as mentioned, they can represent a failure to attempt to mark past tense. Second, the child, while attempting to mark tense, may retrieve the information that an irregular past form of a given verb exists, but fail to retrieve its phonetic content (see Note 7); the past feature or pointer would block the regularization but the child would not have the irregular form at hand either. Third, a child may fail both to retrieve the irregular and to block the regular, and might thus feed the stem into the regular suffixation process, which ordinarily results in an overregularization. But the regular suffixation process does not succeed 100% of the time for regular verbs, so it may occasionally fail here too, yielding the unchanged stem. Finally, some irregular verbs actually show no change between stem and past (e.g., *hit, cut, put, set*), and the child might analogize the no-change pattern to similar irregular verbs.

During the early stages of language development, where stem forms are the great majority of forms uttered (Cazden, 1968; Brown, 1973; see Chapter 12), it is reasonable to assume that the absence of intention to mark is the usual or exclusive cause. Although tense and agreement are marked obligatorily in languages like English, the child cannot be born knowing this. Thus it is not surprising that there might be an initial period in which children have not yet developed any mechanism for systematic marking of any specific inflectional feature. Indeed, bare stems constitute the great majority of young children's verb forms in English, not only in past tense contexts, but in progressive and third person singular contexts (Brown, 1973). Thus the simplest hypothesis is that most of these early stems represent failures to mark tense at all, not attempts that fail for the three possible reasons listed above.

During the later stages, when the child frequently marks tense on irregular verbs, it is plausible that some of the stem errors represent attempts to mark tense that end up unsuccessful. Luckily, we shall see in Chapter 12 that during the time when overregularization errors occur, stem errors become rare, and so little hinges on whether or not one looks at an index that includes them. Where relevant, we discuss these possible effects.

Estimates of No-Marking Errors (Stems in Obligatory Past Tense Contexts)

Though no-marking errors are not counted in the overregularization rate, we will require them in a number of other calculations, such as how often the past tense is overtly marked at all, or how successful regular past tense marking is. Unfortunately automatic computer searches cannot discriminate stems used in obligatory past tense contexts from stems used properly as infinitives, imperatives, and present tense forms. Thus the relevant data must be tabulated by hand by a linguistically sophisticated scorer. This heroic task was beyond the scope of this study, but has already been done for Adam, Eve, and Sarah by Courtney Cazden (Cazden, 1966, 1968; Brown, Cazden, and de Villiers, 1971; Brown, 1973), summarized in unpublished tables that she and Roger Brown have generously provided us. It has also been done for Abe by Stan Kuczaj, summarized in an appendix to his thesis (Kuczaj, 1976).

For certain analyses in Chapters 10 and 12, we will need to combine data on the form of marking (correct irregular versus overregularized) with data on the presence of marking (unmarked stem versus correct irregular). For Abe this is straightforward because all token counts are provided in a single table in Kuczaj (1976). For Adam, Eve, and Sarah this cannot be done because Cazden did not count overregularizations at all (neither as irregular errors nor as correct irregulars nor as correct regulars), and counted fewer correct irregular past tokens than we did. The reason was that she needed a constant definition of "obligatory past context" across stems and correct forms, and many tokens occurred in contexts that were not unambiguously obligatory for past tense and thus could not be included.[23] In order to combine the data sets we adopted the assumption that overregularizations were not any more likely than correct irregulars to occur in "obligatory" contexts as opposed to "non-obligatory" contexts; the distinction, after all, is defined in terms of the knowledge of the observer, not the child. Therefore we calculated what proportion of total irregular tokens (by our counts) Cazden listed, and multiplied this proportion by the number of our overregularization tokens, yielding an estimate of the number of overregularization tokens occurring in obligatory past tense contexts (i.e., an estimate of the number of overregularizations Cazden would have found if she had looked for them among the contexts in which she counted correct irregulars and unmarked stems). These token estimates could then be meaningfully combined with Cazden's counts of irregular and stem tokens.

Other Errors Related to Overregularization

---

[23]Brown, Cazden, and de Villiers (1971) define obligatory contexts as including "adverbs like *yesterday*, marginal notations [comments by the transcriber], expansions [by adults], continuity of tense, etc."; attempts by the child to imitate adult sentences containing tensed forms were also counted.

Note finally that our study excludes two kinds of errors that are sometimes lumped with overregularization but that are logically distinct from it. In languages with richer inflectional systems than English, children often inflect a stem with an affix that is incorrect for some feature of the word, such as gender or animacy, or incorrect for some feature of the context, such as case or definiteness (Slobin, 1973; Pinker, 1984). Pluralizing mass nouns (e.g., *waters*) or extending the third person singular suffix to other persons or numbers (e.g., *we walks*) would be examples in English. Such errors are best characterized as *underdifferentiation* -- an insensitivity to some systematic distinction relevant to inflection -- rather than overregularization, which involves idiosyncratic lexical exceptions to a systematic inflectional process.

Second, we are excluding overapplications of *irregular* patterns to inappropriate irregular verbs such as *tooken* or *brang*. Because irregular patterns, even when used productively, have qualitatively different properties than the regular suffix (reviewed in Sections 14 and 14), they are logically distinct from overregularizations and should be tallied separately. Theories differ as to whether regular and irregular patterns are in fact handled by the same kind of mechanism; according to the particular hypothesis we are considering in this monograph, they would represent competition between two irregular forms for one memory slot, not a competition between a stored irregular and the regular rule (see Ullman & Pinker, 1990, 1991, for discussion). These issues can be sidestepped in the present investigation; Bybee and Slobin (1982a) point out that over-irregularizations are extremely rare in preschool children's language, and there are very few in our samples. For Adam, Eve, and Sarah, Cazden (1966) reports only *beat-bate* and *hit-heet* for Adam, and *beat-bet* for Sarah, to which we can add *bite-bat* for Adam. *Sweepened* was the only such form we happened to come across for Abe, though because there is no mechanical way of searching for such errors, there may be others. Overextensions of irregul. · patterns to regular verbs appear to be even rarer; we are aware only of *trick-truck* from Adam.

10:
The Rate of Overregularization

In this chapter we estimate children's typical overregularization rate in an attempt to resolve the empirical inconsistencies in the literature and to test possible explanations of the nature of overregularization.

As discussed in Chapter 8, the most frequent characterization in the literature is that once children begin overregularizing they do so all the time, replacing correct irregulars altogether. Other characterizations have children varying freely between overregularized and correct irregular forms, either at a rate near the rate of tense-marking regular verbs (if irregulars are not being discriminated from regulars) or at a rate that is not markedly far from values interpretable as representing system-wide indifference (if neither the irregular nor the overregularized form of irregular verbs are systematically favored, but a panoply of factors settles a competition between them for each verb). In either case a preference for the irregular would emerge slowly before culminating in adult performance. Such a stage of free variation would call into question any kind of Blocking or Uniqueness principle (Maratsos, 1987), and would leave unsolved the learnability problem of how children eliminate the incorrect forms.

The third possible empirical pattern is that overregularization errors are rare relative to correct irregulars. If so, there would be no qualitative difference between children and adults. Both would discriminate between regularized and irregular forms, presumably because application of their regularization mechanism to listed irregulars is blocked. Children's occasional overregularizations, like adults' speech errors and adults' uncertainty about low-frequency irregulars like *smote*, could be attributed to probabilistic imperfect retrieval from rote memory. Strictly speaking, it would suffice to show that the rate of regularizing irregular verbs is reliably less than 50% to demonstrate that children possess some mechanism that acts to give the irregular past tense form of a verb priority over the regularized form, contrary to the predictions of all hypotheses other than Blocking. The claim that that mechanism in fact is Blocking (which actively suppresses overregularization), as opposed to some weak statistical preference for the irregular for some other undiscovered reason, would be more convincing the lower the overregularization rate is, the more pervasive such a low rate turns out to be, and the more clearly it can be shown that irregularity per se, not some variable confounded with it, is associated with low rates of using the regular suffix.

Of course there is no single quantity that we can call "Children's Overregularization Rate." Any estimate must aggregate over children, ages, and verbs, and the estimates will necessarily vary depending on which subsets one chooses to include. Therefore after calculating some initial estimates of the overall rate, we will break down the data in a variety of ways, to ensure against various kinds of averaging artifact and sampling error. Ideally, the range of different estimates should not stray too far from one of the predicted values (near 100%, near the regular tense-marking rate, near 50%, near 0%) as one focuses on subsamples.

Overall Overregularization Rate

We first calculated overall overregularization rates. The median overregularization rate across the 25 children with individual transcripts was 2.5%. Table 2 shows the relevant data for the 25 individual children, whose distribution of overregularization rates is plotted in the histogram in Figure 1. The distribution is roughly exponential, with most children at the extreme low end; only two children, April, and Abe, overregularized more than 10% of the time (13.0%, and 24.0%, respectively). The average overregularization rate across the 25 children was 4.2%. The three group databases replicate this figure, with rates ranging from 1.0% to 6.5% and a mean of 3.7%. In Section 10 and Chapter 12, after we have examined some of the factors that affect overregularization rates, we will be in a better position to speculate on why Abe's rate was so much higher than those of the other children; some of the difference, we shall see, may be artifactual. But even he was very far from overregularizing 100% or even 50% of the time, and his is the highest rate we see. Thus, the global data suggest that overregularization is a relatively rare phenomenon; if they legitimately reflect children's tendencies, it would suggest that children's language systems, like adults', are strongly biased to suppress overregularization, contrary to common belief.

Figure 1 and Table 2 here

Change in Overregularization Rate Over Time

Of course, the surprisingly low rates obtained may be an averaging artifact: each child could go through a circumscribed U-shaped period of constant or indiscriminate overregularization, preceded and followed by many more months of near-perfect performance. For example, if there were a span in which overregularization did not occur for 28 months, but went up to 96% for one month, the average overregularization rate for the span would be only 3.3%. This possibility is shown in Figure 2. (All our developmental graphs plot the proportion of past tense forms that are correct (100 - overregularization rate) rather than the proportion overregularized, so that regressions in development appear as U's, not inverted U's.)

Figures 3-6 plot monthly overregularization rates (pooled across 2 to 4 transcripts) for the children whose overregularization rates we tabulated longitudinally (Adam, Eve, Sarah, and Abe); the data are provided in Appendices -. For all four children overregularization begins early in the span sampled and lasts for the rest of the period. Aside from Eve, whose samples end at 2;3, we see fairly steady overregularization from some time in the third year at least into the sixth year. Only for Abe is there any hint of a reduction in overregularization rate late in development.

The graphs demonstrate that low rates characterize the entire period of overregularization. Adam's highest monthly rate of overregularizing is only 6.8%; Eve's is 23.1%; Sarah's is 15.8%. Even Abe, an outlier among the 25 children, displayed an overregularization rate of only 47.6% in his most extreme month (see Figure 6); this is the highest rate we find among 109 monthly estimates from the four children who were examined longitudinally.[24]

Figures 3-6 here

Is there a U-Shaped Developmental Sequence?

The ubiquitous claim that children pass through a U-shaped developmental sequence in acquiring irregular verbs (in particular, that their performance declines at some point) is based on observations by Ervin and Miller (1963), Miller and Ervin (1964), and Cazden (1968) that a few irregular forms were often used correctly by the children before they started to overregularize. U-shaped development has never been documented quantitatively, however, and Stemberger (1989) and Marchman (1988) have questioned whether it exists. In this subsection we define the phenomenon precisely and test for it in our data.

Many measures of children's performance that one plots against time will show some kind of dip, for a variety of reasons, and many others will not. Until one specifies with precision what a "U-shaped" sequence is supposed to refer to, one cannot test whether such a sequence exists or what causes it. We will follow the definitions of Cazden (1968) and

---

[24]Note that it would not even be sound to conclude that Abe went through a stage at which his overregularization rate was 47.6%, because of the statistical phenomenon of regression to the mean. This particular month was chosen post hoc *because* of its high rate. Any monthly estimate reflects a sum of the child's true overregularization rate (a psychological tendency that could be measured as how frequently the child would overregularize over a very large number of comparable opportunities at that age) and an error component (from sampling and other sources of noise independent of the true rate). By deliberately selecting the month with the maximum observed overregularization rate, we are simultaneously selecting for samples with high true overregularization rates and for samples with high noise in the direction of overregularization. The true rate, therefore, would be expected to be lower than the sampled rate in such months. For similar reasons it would not be sound to point to Abe's data from 4;8 and conclude that he went through a stage in which his overregularization rate was zero.

Rumelhart and McClelland (1986), who were referring to a transition from a period in which past tense forms are marked correctly whenever they are marked at all to a period in which some overregularization errors occur as well (see Section 9 of Chapter 9). The previous section showed that when overregularizations appear, they neither predominate, nor alternate with the correct irregulars, but are always a minority. However, one can still test whether a period of extremely accurate performance precedes the first overregularizations; in other words, whether children at some point in development get worse (even if they never get very bad).

Our sample includes 9 children with extended longitudinal data. For two of them, Abe and Nathaniel, the first transcript contains an overregularization. If these transcripts happen to coincide exactly with the boys' very first uses of irregular past tense forms, the transcripts show that they did not undergo U-shaped development. But it is more likely that the transcripts began well after irregulars of one or both kinds were being produced, in which case the transcripts are uninformative. Note that for the same reason, any set of transcripts that does not begin early enough in the child's language development is apt to underestimate the length of any overregularization-free sequence in which irregular pasts are used.

For the other seven children, one does see correct irregulars in the transcripts before the appearance of the first overregularization (not surprisingly, since three of them -- Adam, Eve, and Sarah -- were the ones studied by Cazden). However, the impression of a developmental change could be a sampling artifact. Since we now know that children's overall overregularization rate is low, their tendency to overregularize could be unchanged throughout development, but their early samples might simply be too small to contain any examples of overregularization. Imagine drawing playing cards from a deck with replacement, looking for a black king (whose frequency in the deck is approximately equal to children's overregularization rate). One might have to draw a large number of cards before the first one appears, even if the deck is complete and properly shuffled.

It is not legitimate to test for a change in rate by comparing overregularization rates before and after the first overregularization, because the post hoc nature of the dividing line will inflate the chances of obtaining a spurious difference. One stringent test can be done as follows. If the child's overregularization rate is p, then under the null hypothesis of no change in this rate over time, the chances that the first irregular verb form in the sample will be correct is $(1 - p)$. If the likelihood of a child overregularizing an utterance is unaffected by whether or not the child overregularized the previous past tense utterance (an assumption we will examine below), then the chances that both of the first two utterances will be correct is $(1 - p)^2$, the chances that the first three will be correct is $(1 - p)^3$, and so on. One can test whether there is an improbable run of consecutive correct irregular past tenses at the beginning of a child's records by calculating $(1 - p)^n$, where n is the number of irregulars in the transcripts preceding the transcript containing the first overregularization. (Correct forms preceding the first overregularization within that transcript would be excluded, because including them would correspond to the rather implausible assumption that the child had brought the regular rule on-line for the first time in the midst of that very session.)

Table 3 shows the results. For Adam, Sarah, and Peter, the probability of obtaining a string of correct utterances of the observed length or less if the early overregularization rate was the same as that for the entire corpus is very small (.0012 or less). For the other children considered individually, no conclusions about U-shaped development can be drawn: The first overregularization, though coming after a number of correct forms, does not appear substantially before it is expected to given the children's overall overregularization rates. However the seven children can be treated as a single sample in a meta-analysis, allowing us to assess the aggregate probability of observing such longer-than-expected strings of correct forms under the null hypothesis. Following Rosenthal (1984), we converted the children's individual probabilities to z-scores, summed them, divided by the square root of the number of children (7), and transformed the resulting *z*-score back to probability values; pooling the children's probabilities, not their sentences, avoids an obvious averaging artifact. The overall probability of obtaining the data under the hypothesis of a constant overregularization rate is .000045.

Table 3

These analyses crucially depend on the assumption that the probability of overregularizing is independent of whether the preceding irregular past tense utterance was correct or an overregularization. We tested the validity of the analyses in two ways. First, we estimated the validity of the assumption directly on a sample of data, namely Abe from 3;3 to 3;4. His conditional probability of producing a correct irregular past (as opposed to an overregularization) given that his preceding irregular past tense form was correct was .58; his conditional probability of producing a correct irregular past given that his preceding irregular past tense form was an overregularization was .62. The difference is negligible and not in the direction that would compromise the probability analyses. Second, we tested the robustness of the conclusions under a pessimistic assumption of one kind of violation of independence: that every other correct verb form in a pre-overregularization string was completely determined by a perseveration effect from the preceding correct verb form. Under this assumption each string of correct forms would be effectively half as long originally measured, and so we divided the number of initial consecutive error-free pasts in Table 3 by two. The probabilities of obtaining strings of correct forms of those lengths or greater, under the hypothesis of no change in underlying overregularization rates, remain less than .05 for Adam, Peter, and Sarah individually, and less than .06 for the aggregate sample in the meta-analysis.

In sum, there is quantitative evidence that children's first overregularization follows an extended period in which their overtly tensed irregular verbs are all correct, and this effect can be demonstrated very strongly for some individual children. In this sense children do get systematically worse as they get older.

Other senses of "U-shaped development." Some discussions in the literature either claim that there is no such thing as U-shaped development of irregular tense marking, or that connectionist models developed after Rumelhart and McClelland's can account for it better than their model did. These discussions, however, refer to very different phenomena than the one discussed by Ervin and Miller, Cazden, and Rumelhart and McClelland. We review these other senses of "U-shaped development" and point out the extent to which the corresponding empirical claim is supported in our data.

In the network model described by Plunkett and Marchman (1990), early-acquired verbs were permanently resistant to overregularization; the so-called "onset of overregularization" in their model pertained to its performance on newly-acquired verbs.[25] In actuality, once children begin to overregularize, they produce errors for many of the verbs that earlier they had produced correctly (that is, the sequence does not consist of correct performance for some irregulars early, and overregularization only for newly acquired verbs, with the early correct ones eternally protected.) For Adam, 15 of his 23 overregularized types (65%) had been produced correctly at least once beforehand. For Eve and Sarah, the respective figures were 3/9 (33%) and 15 out of 26 (58%).

Stemberger (1989) suggested that Rumelhart and McClelland (1986) were misled by the psycholinguistics literature into trying to model a nonexistent phenomenon. He plotted data from Abe showing no signs of U-shaped development. However Stemberger is simply discussing a different quantity than the one Ervin and Miller, Cazden, and Rumelhart and McClelland focused on. What he plotted was not overregularization rate, but a different measure, overall correct performance, defined as 1.0 minus

[25]"Interestingly, however, the initial 20 verbs are likely to continue to be mapped correctly by the network, even in the presence of the other erroneous mappings." (Plunkett and Marchman, 1990, Section 4, p. 32).

$$\frac{\text{\#overregularization tokens + \#stems in past contexts}}{\text{\#overregularization tokens + \#stems in past contexts + \#correct irregular past tense tokens}}$$

The data in Appendices - confirm that the overall rate of correct performance does not decline notably with age for any of the children. But it is not clear that anyone in the psycholinguistics literature has ever claimed it did, so the point of Stemberger's objection is unclear.

Plunkett and Marchman (1990, 1991) refer to yet another sense of the term "U-shaped development" in arguing for the psychological reality of their connectionist models. In Plunkett and Marchman (1991), the learning curves all start out at levels of performance far *less* than 100%, and then increase; the authors call the small wiggles in this overall increasing curve "U-shaped development." Although all the children we examine show local ups and downs in their monthly measures of overregularization rates, there are many explanations of these blips, of which sampling error is the simplest.[26]

Finally, in Plunkett and Marchman (1990), any verb that is used once correctly and then once incorrectly is characterized as undergoing U-shaped development, which is misleading for a different reason: Any stationary stochastic process (e.g., a string of coin flips) will produce local sequences with such patterns. Neither the fact that children produce such sequences, nor the ability of their model to do so, is surprising.

When Does Overregularization Cease?

There are no signs of overregularization going away or even decreasing in Adam,

---

[26]Kruschke (1990) makes the point a slightly different way:

> The recent work of Plunkett and Marchman [1991] does *not* exhibit U-shaped learning, contrary to their claims. They showed that acquisition fluctuated depending on the particular training sequence, but they failed to mention that *on average* their model showed monotonic, not U-shaped, acquisition. (p. 61)

Eve, or Sarah's samples, which last through the early 5's for the former and latter. Perhaps because Abe's overregularization rate is higher to begin with, one can discern a trend of gradual overall improvement (the slope of the best-fitting straight line is +6.0 percentage points per year), superimposed on seemingly random fluctuations that leave many late months with higher overregularization rates than early ones. Clearly, overregularization diminishes extremely gradually. Kuczaj (1977a), based on a personal communication from Slobin, notes that it is still present in school-age children of 9 or 10.

Two studies provide us with estimates of the overregularization rate in older children. Moe, Hopkins, and Rush (1987) report a sample of 10,530 irregular past tense utterances from over 329 first graders. The overregularization rate in their data is 2.8%.

Carlton (1947) reports 2196 past tense tokens among the speech of 96 fourth-graders she recorded. These included thirteen overregularizations. If approximately 75% of past tense tokens in speech are irregular (see Chapter 11), the overregularization rate for this group is 0.8%.

Do overregularization errors ever completely disappear? Joseph Stemberger has kindly provided us with the full set of past tense overregularizations in his corpus of 7,500 adult speech errors. The list includes 25 past tense overregularizations (18 of the stem, 7 of the irregular past form); Stemberger (1989) suggests that the rate of adult speech errors might be about 1 error per 1,000 sentences. If we assume that all sentences contain verbs, that about 10% of verbs in casual speech are in the past tense (Smith, 1935; Adams, 1938), and that 75% of adults' verb tokens are irregular (Slobin, 1971; see Chapter 11) we get a very crude estimate of adults' overregularization rate of .00004 -- three orders of magnitude lower than preschoolers', and two orders lower than 4th graders.

So although in one sense both children and adults overregularize, there is also a dramatic difference in their rates of doing so. Perhaps the difference is just a consequence of hearing more tokens of each irregular verb as one lives longer, with more exposures leading to more reliably accessible memory traces. For example, a negative exponential learning curve with a time constant of one order of magnitude of improvement in retrieval probability for every 5 years' worth of irregular past tense tokens could handle the reported overregularization rates from the preschool years through adulthood comfortably. In the absence of more plentiful and finer-grained data, it is premature to claim that there is no qualitative difference between children and adults, but current evidence does not demand that there be a difference.

Differences in Overregularization Rate Among Verbs

Another possibly misleading effect of averaging could result from combining data on individual verbs: Perhaps a few very commonly used verbs are never or rarely overregularized, but most verbs are overregularized most of the time or at least indiscriminately. Because these verbs may differ from one child to another, the question can only be addressed by examining overregularization rates for different verbs in individual children. The relevant data for Adam, Eve, Sarah, Abe, and the other children are presented in Appendices -. Aggregate measures of overregularization rate, described in Chapter 13, are presented in Appendix .

It is important to note, however, that sampling error can make estimates of overregularization rates for particular verbs of particular children extremely misleading. In the extreme case, if a verb is only used once, its observed overregularization rate can only be 0% or 100%, regardless of the child's actual overregularization tendency. If it is used twice, the observed rate can be 0%, 50%, or 100%, and so on. Even with a low overall regularization rate and slightly larger samples, high estimates of overregularization rates for a given verb frequently will arise by chance. Histograms of the overregularization rates for all of a child's verbs show spikes at values corresponding to ratios of small integers such as 0, 1/4, 1/2, and 3/4, and 1. But if we restrict attention to only those irregular verbs that were used 10 times or more in the past tense (the cutoff suggested in Chapter 9), we find that

Adam did not overregularize any of his 32 verbs at a rate higher than 10%, Eve overregularized *fall* 8 out of 10 times but did not overregularize any of the other 9 more than 20%, and Sarah overregularized *throw* 7 out of 10 times but did not overregularize any of the other 26 more than 33%. Histograms of Adam, Eve, and Sarah's overregularization rates for verbs used a minimum of ten times are displayed in Figures 7-9.

Figures 7-9 here

Even for Abe, the extreme overregularizer, most commonly used verbs fall at the low end of the distribution of overregularization rates (mode 10-20%, median 30%, mean 32%). This is shown in Figure 10, a histogram of overregularization rates for Abe's verbs used 10 times or more.

Figure 10 here

It is clear that the overall low overregularization rate is not an artifact of averaging a few verbs that are never overregularized but used extremely frequently with a majority of verbs that are usually overregularized but used less often. Rather, low overregularization rates characterize most of children's commonly-used verbs. Nonetheless, it is also quite clear (especially for Abe) that different verbs are overregularized at different rates, and estimates of overall overregularization rates must be qualified by examining different verbs. For Adam, Eve, Sarah, and Abe, Chi-square tests amply show that overregularization rates differ among verbs used more than 10 times (all *p*s < .001). It is statistically permissible to perform such Chi-square tests for four other children in our sample (GAT, Naomi, Nathan, and Peter); the tests are significant for all of them. Chapter 13 is devoted to investigating properties of verbs that make them more or less likely to be overregularized. The most important, not surprisingly, is frequency: verbs that parents use in the past tense less often are more likely to be overregularized by their children.

It is important to note that the correlation between verb overregularization rate and verb frequency necessarily interacts with any cutoff designed to eliminate small samples. The verbs that parents use less often are also the verbs their children use less often (see Chapter 13). Therefore excluding verbs likely to give rise to sampling errors will also tend to exclude verbs with higher overregularization rates. Appendices - confirm that many infrequently used verbs yield high estimated overregularization rates. In fact computing the unweighted mean of the overregularization rates of all verb types yields figures of 22% for Adam, 17% for Eve, 18% for Sarah, and 47% for Abe; the mean of the type means for all 25 children is 10.5%, the median, 11%. One might wonder whether the relatively high type mean for Abe when it is calculated over all of his verb types casts doubt on the conclusion that his grammatical system suppresses overregularization. Such doubts are not warranted. Means calculated over types do not reflect the functioning of the child's grammatical system, because a verb used once is weighted exactly as strongly as a verb used a thousand times. Rather, what the typewise mean is capturing is the distribution of frequencies of the child's irregular verb vocabulary: the more low-frequency verbs a child uses, the higher the mean. In fact, the theory invoking blocking and retrieval failure, though it predicts low overregularization rates *overall*, also predicts *high* overregularization rates for low-frequency verbs: 100%, for verbs that have never been heard or attended to in the past tense form (like *shend* for adults), lower but still high rates for verbs that have been heard occasionally, and so on (see Chapter 8). In a sense, a child's seldom-used verbs are misleading for two reasons: we cannot be confident about the actual overregularization rate of an irregular verb used once or twice, and a child cannot be confident that a verb is irregular if he or she had only heard it used once or twice.

Irregular verbs are, quite generally, overregularized at low rates, but these irregulars can be very different from one another. This suggests that the overall low overregularization rate reflects some phenomenon that suppresses overregularization globally across the irregulars (modulo frequency), and is unlikely to be attributable to any general bias of the child to favor correct irregulars because of some lexical property that many of them happen to possess. For example, it might be argued that children's grammatical systems allowed both overregularizations and correct irregulars, but that they had a response bias in favor or producing irregulars because they tend to be shorter and simpler than overregularizations, and this is the cause of the measured overregularization rate being less than 50%. Such a counterexplanation can easily be ruled out. Among the irregular verbs our subjects used, 14 have correct past tense forms that are the same length or shorter than their overregularized counterparts, usually with a similar phonological structure: *be, buy, do, feel, go, hear, keep, leave, lose, mean, say, sleep, sweep*, and *tell*. If simplicity, not irregularity, were responsible for the overall low overregularization rates, then these verbs, lacking that advantage, should show overregularization rates much closer to 50%. Instead, the overregularization rates for tokens of these verbs were 0.7% for Adam, 9.3% for Eve, 2.8% for Sarah, and 19.9% for Abe; the mean overregularization rates across these verb types were 18.3% for Adam, 5.25% for Eve, 5.0% for Sarah, 36.9% for Abe, and about 8% for the aggregate rates from all 19 children who overregularized at least once (see Chapter 13); medians were far lower, usually 0%. All of these figures are close to or lower than those obtained for the full set of irregular verbs. More generally, the data suggest that the only property of irregular past tense forms that is likely to account for children's strong tendency to prefer them to overregularizations is irregularity itself.

Changes in Overregularization Rates Over Time among Different Verbs

The most stringent test of the hypothesis that overregularization is a probabilistic and relatively rare event would look at the fate of individual irregular verbs for individual children as they grow older. This is independent of the overall level of overregularization for different verbs that we have just examined, just as the waveform of a sound wave is independent of its amplitude and its DC component. For example, it is possible that each child goes through a stage for each verb during which the verb is overregularized exclusively (see Figure 11), or as often as it is produced correctly. If these stages are fairly brief and circumscribed, the steady low rate of regularization could be an averaging artifact of a sequence of deep narrow U's, one for each verb. If so, or if the verbs all follow some other set of out-of-phase developmental curves, the protracted period of overregularization would reflect a failure to apply blocking to different verbs at different times.

Figure 11 here

There are other possible interactions among children, verb, and age that would be noteworthy. One can determine whether any verbs cease to be overregularized altogether before the end of the period, or alternatively, whether a child begins to stop overregularizing all verbs at the same time; the latter finding would suggest that the child learns or develops the Blocking principle only at that point. Another possibility is that various verbs follow different and largely unsystematic patterns, perhaps because, as we have suggested, overregularization is a quasirandom performance deficit.

Clearly, low rates of overregularization are not an artifact of a sequence of transient overregularization stages, one for each verb: Individual verbs can be overregularized across large spans of time. The very first verb that Adam overregularized was *feel* at 2;11; he also overregularized *feel* in his last sample at 5;2. Similarly, *throw* was overregularized at 3;4 and at 4;4; *make* at 3;5, and 5;2; *fall* at 3;5 and 4;10. Sarah's first overregularization, *heared* at 2;10, appeared again at 4;11; *winned* and *maked* also made appearances in the samples separated by a year or more. Even the 9 months' worth of samples from Eve contain *falled* at 1;10 and again at 2;2. More than half of Abe's overregularized verbs (44) were overregularized over a span of one year or more; 32 were overregularized over a span of two or more years.

Unfortunately, when we turn away from the simple question of whether overregularizations of a given word reappear across long time spans and try to trace each one of a child's irregular verbs over time, we run up against severe sampling limitations. As mentioned, small samples yield inaccurate estimates of overregularization rates, and many of the samples of tokens of a given verb for a given child in a given month were very small. Thus, developmental curves for individual verbs with low token frequencies for the child can oscillate wildly among a few discrete values, revealing little about changes in the underlying true rates. With these caveats in mind, we now examine curves for individual verbs for Abe, the most prolific overregularizer.

For most of the 70 irregular verbs Abe used, the curves can best be described as chaotic, and highly variable from verb to verb. They are most conveniently summarized by pointing out four rough patterns, shown in Figures 12 through 15. Some verbs, like *eat* (Figure 12), are overregularized in the earlier transcripts but appear to be completely mastered before the end of the sampling period. (Other verbs with this pattern include *cut, fall, go, make, think,* and *throw.*) A second class of verbs, such as *say* (Figure 13) are rarely overregularized at any point. (*Find, forget, see*, and *tell* are others.) A few, such as *draw* (Figure 14) are overregularized throughout the sample (*build* is similar); such verbs were used only rarely, and might be overregularized because they are also rare in parental speech, though sampling error cannot be ruled out. But many verbs, such as *win* (Figure 15), *bite, break, blow, buy, catch, come, feel, get, know, put*, and *shoot*, show no interpretable trend, oscillating between samples with high and low measured overregularization.

Figures 12-15 here

In sum, apart from haphazard variation (possibly due to sampling error) and overall low or high rates (to be discussed in Chapter 13), the only meaningful temporal pattern for individual verbs seen in Abe's data is that a few appear to be mastered in the late 4's, thereafter resisting overregularization completely.

How High Can Estimates of Children's Overregularization Rate Go?

Given that the overregularization rate varies across verbs, and to a lesser extent, children and ages, there is no absolute "overregularization rate"; the estimates will change with methodological decisions of which verbs, children, and ages to include. The low estimates in Section 10 included all the samples meeting the criteria described in Chapter 9, without any exclusions based on how the data came out. Still, it is conceivable that these criteria adventitiously included some samples of questionable representativeness, relevance, or accuracy. In this section we will test the robustness of the low estimated overregularization rate, by eliminating all samples that could be challenged, for any plausible reason, as potentially deflating the measured overregularization rates. If these worst-case estimates continue to be low after steps are taken to bias them as high as possible, one can have more confidence that there is some real process causing them to be low and that the effect is not an artifact of the adoption of one selection criterion or another.

First, for some of the children the samples are small, and we may have caught them in conversations where few irregular verbs were used in the past tense, or perhaps only a small number of well-mastered ones. Hence we can eliminate all children who contributed less than 100 irregular verb tokens to the overall sample.

Second, because of possible U-shaped developmental sequences, some of the youngest children sampled may not have entered the stage at which overregularization had begun. If so, it would be misleading to count their samples from the pre-overregularization period (with rates of 0%, by definition), because the psychological process of interest was not yet occurring. Moreover even if overregularization had b[illegible] may not have reached its peak. So, following the suggestion of one of the referees [illegible]arlier draft, one might question the inclusion of Allison, April, Nat, and Peter, none of whom were recorded any later than 3;1. These children could not have been excluded [illegible]ause Abe had his *worst* month of overregularizing at 2;10, and Eve overregularized at [illegible],8. But let us assume that these four children were on the verge of overregularizing at high rates, and exclude their data. We will continue to include Eve, even though she falls within the age range of the children excluded here for being too young; this decision renders overregularization estimates higher than they would be if she were excluded. The rationale might be that her language development was precocious and hence her true overregularization tendency was indeed manifested in the available samples.

Third, we might be sampling overregularization-shy children from the opposite arm of the U, where the errors have already disappeared. The Hall, et al. children might be suspect for these reasons. (In fact, data from the children we analyzed in longitudinal detail -- Adam, Sarah, and Abe -- do not themselves call for such an exclusion, because while Abe showed an overall decline in in the age range corresponding to the Hall, et al., children, he was still overregularizing at rates from 0 to 22%, and Adam's and Sarah's overregularization rates *increased*.)

Turning now to the different samples within longitudinal records, we can consider the fact that there was a demonstrated overregularization-free period in the early months of Adam, Peter, and Sarah. So here we will exclude all correct forms in the transcripts preceding the transcript in which the first overregularization is found.

Fifth, perhaps not all verbs deserve to be included. *Get* could be problematic for children other than Abe, for whom Kuczaj (1976) carefully coded all forms as the past of *get* versus the present of *got*. For Adam, Eve, and Sarah, recall that we checked contexts and excluded present tense forms, but some cases counted as past forms could have been present forms. For the other children the computer counted all instances. We will report estimates that exclude all instances of *got* for children other than Abe.

Sixth, one might argue that *have*, *be*, and *do* (included only for Adam, Eve, and Sarah) are special. Even though we did not include any auxiliaries in our analyses, these verbs obviously share their morphemes with the auxiliaries, and are irregular not only in the past but in the third person singular present. Conceivably the child might treat non-auxiliary versions of *have*, *be*, and *do* similarly to the auxiliary versions, and differently from other verbs. Actually, when Stromswold (1990) examined overregularization rates for auxiliary and non-auxiliary forms of these verbs in a large sample of ChiLDES children overlapping with the one we use here, she found that the non-auxiliary versions -- that is, the ones we have been counting -- were overregularized at rates comparable to those of the other verbs we report here. Stromswold, counting correct forms from the children who made overregularization errors during the developmental spans in which they made them (a procedure that yields slightly higher estimates than the ones we report) obtained overregularization estimates of 2.5% for *have*, 4.3% for *be*, and 7.5% for *do*); see also Appendices -.) But again we will report the effects of excluding them.

Table 4 here

Table 4 summarizes the results of various exclusions on the estimated overregularization rates. The results are clear. Even combining all these exclusions in an effort to avoid every possible bias toward low estimated rates, we are only able to push the estimate up to a mean across children of 9.8% and a median across children of 8.7%, ranging from 3.6% to Abe's 24.0% (which we will focus on in section 10). These worst-case estimates are still far closer to 0% than they are to 100%, 50%, or, for the bulk of the children, even 25%. They are low enough that it seems difficult to avoid the conclusion that the child's grammatical system contains some mechanism that, while allowing overregularizations to occur, is strongly biased against them in favor of the correct irregular counterpart.[27]

Comparison of Overregularization with Regularization

---

[27]Moreover, the shape of the developmental curves for Adam, Eve, and Sarah are virtually unchanged, and the tests of U-shaped development to be reported in Section 10 all remain significant.

The fact that the absolute rate of overregularization is quite low does not disprove the suggestion (e.g., in Clark, 1987) that children fail to discriminate regular from irregular verbs during the period in which they are overregularizing. Conceivably children inflect *regular* verbs only 2.5% of the time during this period, leaving the regulars uninflected the remaining 97.5% of the time. If so, irregular past forms would not be blocking the regularization of irregular verbs, as the blocking-and-retrieval-failure hypothesis claims; instead, the low rate of overregularization would presumably mean that regular suffixation was an extremely errorprone process, failing most of the times it is invoked. Correct irregular forms would continue to be used, together with stems, in the circumstances in which the regularization process failed.

The data necessary to evaluate this suggestion consist of estimates of the rates of inflecting regular verbs in obligatory past tense contexts. These estimates are presented in full in Appendices -. They clearly show that children are not failing to discriminate regular from irregular verbs. The rates of correctly affixing regular verbs in obligatory contexts in the months including and following the first recorded overregularization, summing over tokens, are 73% for Adam, 66% for Eve, 85% for Sarah, and 97% for Abe. These are way higher than the children's overregularization rates, and the difference can be seen in all 65 individual months for which data are available for the four children (for Adam, Eve, and Sarah, these are the months for which Cazden coded proportion of use in obligatory contexts, which do not extend to the end of the transcripts). More pertinently, we can compare the rate of marking tense on regular verbs in obligatory verbs with the rate of producing overregularizations as a proportion of all obligatory contexts, or

$$\frac{\text{\#overregularization tokens (in obligatory contexts)}}{\text{\#overregularization tokens + \#irregular past tokens + \#stem tokens (in obligatory contexts)}}$$

where our counts of the number of overregularization tokens are scaled to make them commensurable with Cazden's tallies, using the procedure explained in Chapter 9. These overregularization rates were 0.6% for Adam, 4.5% for Eve, 1.1% for Sarah, and 23.2% for Abe, far lower than the rate of regular marking. Moreover these rates were lower in all 65 individual months.

Though children clearly discriminate irregular from regular verbs in general, it is not clear that they discriminate them in the circumstances that lead to overregularization errors. We have suggested that overregularization is caused by a failure to retrieve an irregular past tense form, leading to the application of the regular rule to the stem. If retrieval is all-or-none, then at the moment the child attempts to retrieve the irregular form but fails, the verb should be indistinguishable to the child from a regular form, and should be overregularized by the same process that supplied the affix of a correct regular past tense form. Not all retrieval failures should lead to overregularizations, just as not all regular verbs in obligatory past tense contexts are successfully marked for past tense. But among usages of irregular verbs that are not correct irregular past forms, the proportion that are overregularizations (as opposed to being left unmarked), which equals

$$\frac{\text{\#overregularization tokens (in obligatory contexts)}}{\text{\#overregularization tokens + \#stem tokens (in obligatory past contexts)}}$$

should be identical to the proportion of regular verbs that are successfully affixed (as opposed to being left unmarked) in past tense contexts.

Stemberger (1989) compared these proportions on two of Abe's samples, and found that the overregularization rate calculated as a proportion of non-irregular past tense attempts (overregularizations plus stems) was far lower than the rate of inflecting regular verbs correctly. (The rest of Abe's samples contain very few unmarked verbs in past tense contexts, either regular or irregular, so both marking rates are close to 100%; see Appendix). We performed an analogous test on Adam, Eve, and Sarah, again after adjusting the number of overregularizations to be commensurable with Cazden's counts. Replicating Stemberger, we find that the overall rates of overregularization given non-use of the irregular during the children's periods of overregularization (13.3% Adam, 15.9% for Eve, 10.4% for Sarah), were less than their rates of inflecting regular verbs (66%-97%), and that this was true for 32 of the 33 individual months available.

This analysis shows that when children fail to use an irregular past tense form in a past tense context, they do not treat it identically to a regular verb; they overregularize the irregular verb far less often than they correctly suffix a regular verb. In other words, irregular verbs tend to suppress application of the regular process even when they are not supplied in the correct irregular past tense form. Assuming that there is no consistent reason why children should decide to mark tense on irregular verbs less often than on regular ones, there are two plausible explanations for this difference. One is that the child is analogizing the no-change pattern seen in verbs like *cut* and *set* to other irregular verbs. Later, we review abundant evidence that verbs that end in *t* or *d* are particularly susceptible to this interference. Second, as mentioned in Note 7, retrieval of an irregular may not be all or none. On some occasions, the child may not retrieve the content of an irregular entry (or may not retrieve enough of it to allow the form to be articulated), but might retrieve the information that a past tense entry exists (i.e., the child might retrieve only the pointer or tag to the entry). Since it is the existence of a past tense entry that ordinarily blocks regularization (rather than the particular phonetic content of any irregular form), such piecemeal retrieval would be sufficient to prevent an overregularization without actually supplying the correct irregular form.

Comparison with Previous Estimates in Spontaneous Speech

Given the stereotype that children go through a stage in which they always overregularize, or even overregularize in free variation with correct irregular forms, our finding of a consistently low rate of overregularization across children, ages, and commonly used verbs comes as a surprise. Why has the low rate not been noted before?

One reason is that few investigators have counted the number of correct irregular past tense uses; only a list of errors is reported. A second reason is that most summaries in the secondary literature, such as theoretical book chapters and textbooks, do not even try to provide empirical citations to document the overregularization rates they assume when discussing theoretical implications, as the quotations in Section 8 show.

Indeed, one source that does, Maratsos (1983), provides estimates that bear no obvious relation to the study he cites. Recall that Maratsos claimed that "overregularization, in general, ranged in frequency from .20 to .60 of the children's uses," based on an unpublished talk by Kuczaj (1977b). Maratsos is referring to the study Kuczaj reported in his thesis Kuczaj (1976) and in a published report (Kuczaj, 1977a), and these sources list a much lower range of overregularization rates calculated over uses: .01 to .39 for different children, and .12 to .24 for different subclasses of verbs.[28] Breaking down the rates according to individual children's performance with different subclasses (many obviously with tiny sample sizes) gives a wider range, but the range runs from 0.00 to 1.00, not from .20 to .60. Maratsos (personal communication) has explained to us that he in fact derived these estimates from Kuczaj's unpublished data by averaging overregularization rates over verb *types* for each month, not "uses," as the paper says; the range of monthly typewise means ran from .16 to .69, with the bulk of the months roughly falling within the .20-.60 range that he reports. As discussed in Section 10, averages over types should be systematically higher than the true overregularization rate because they overweight the rare and hence difficult verbs.

Let us now examine the actual data on overregularization rates in preschoolers in the original studies that report them. Bybee and Slobin (1982a) report figures from a pooled sample of 31 children; the rates are listed separately for different subclasses of irregular

---

[28]These figures, taken from Kuczaj's (1977a) tables, refer to overregularizations as a proportion of total usages of irregular verbs in past tense contexts, including unmarked stems. Since most of the children's errors were overregularizations, not stems, the figures for overregularization rates as a proportion of overtly marked verbs are similar, ranging from .01 to .40.

verbs. However, these figures are also averages over verb types, ranging from .1 to .8 for the different subclasses. The numbers of verb types and tokens in each class are not reported, so actual overregularization rates cannot be computed from the data. Fortunately, Slobin (1971) published actual token frequencies of individual verbs summed over 24 of these 31 children (those from the Miller and Ervin, 1964, samples). From his table one can calculate that the pooled children overregularized at a rate of 10.2%, far closer to our estimates.

Recently, Valian (in press) reported overregularization rates in a cross-sectional sample of 21 children, divided into four groups based on Mean Length of Utterance, a measure that roughly correlates with grammatical development in English-speaking children (Brown, 1973). The mean ages of the four groups ranged from 2;0 and 2;7, and the mean MLU's ranged from 1.77 to 4.22, the range spanned by Adam and Sarah between 2;3 and 3;8 and by Eve between 1;6 and 2;3 (Brown, 1973). Valian included the main verb versions of *have*, *be*, and *do*, and excluded their auxiliary versions. Using the individual subject data that Valian has kindly provided to us, we find that children in the first two groups (mean age 2;0 and 2;5) produced 67 correct irregular pasts and no overregularizations. The children in the third group (mean age 2;5, MLU 3.39) produced 109 correct irregular past tense forms and 9 overregularizations, and had a mean overregularization rate of 8.4%. Of the 3 children in the fourth group (mean age 2;7, MLU 4.22), one produced 15 overregularizations of a single verb, but the other 32 irregular past tokens were correct; their mean was 13.2%. The mean overregularization rate of the children in the study was 5.1%.

The ChiLDES archive contains a number of transcript sets that we excluded for the methodological reasons listed in Chapter 9. However it is of some interest to check whether they are consistent with the samples we did report. Among the transcript sets that are from standard American English-speaking children and in CHAT format, there were 3 sets of transcripts from individual children, and 6 group transcripts, ranging in size from 5 to 1596 irregular past tense tokens (median 54). The mean overregularization rate of the three individual children was 12.5% (range 5.1% to 16.7%). The 6 group databases showed overregularization rates ranging from 0% to 10.8%, with a mean of 3.25%.

The only partial anomaly in the empirical literature consists of the data reported by Kuczaj (1976, 1977a). Qualitatively, those data are consistent with our finding of a pervasive bias against overregularizing: neither the mean overregularization rate for the cross-sectional sample, nor that for Abe, come close to the rates of 50% or higher that had been assumed in the literature. But there are quantitative differences. As we have seen, Abe is the outlier from among the 25 English-speaking children we studied in the ChiLDES database. Kuczaj's cross-sectional sample (1977a), of 14 children from 2;6 to 5;6 includes 6 with overregularization rates comparable to those reported here (1.1% to 8.33%), but for the other 8, the rate ranged from 26.1% to 40.2%, with an average over all the children of 20.9%. Thus neither Abe nor most of the 14 children in Kuczaj's cross-sectional sample appear to come from the distribution of unselected children we analyzed (most of whom had very similar backgrounds to Kuczaj's subjects). But Abe and the cross-sectional sample of children appear to come from the same distribution. What Abe and the sample have in common is that they were both studied by Kuczaj, as part of the only investigation we discuss that recorded children's speech expressly for the purpose of studying verbal inflection. Could there be some systematic factor in Kuczaj's methods that would push estimated overregularization rates upward?

According to his thesis, Kuczaj (1976) selected his 14 children from a larger sample of 23, according to the criteria of "(1) clarity of speech, (2) willingness to play games (in particular, to imitate model sentences when asked), and (3) willingness of the parents to aid the investigator by participating as experimenters (that is, by playing games with their child)." Among the ways that the parents aided Kuczaj were supplying the child with erroneous feedback, choosing the times to turn the tape recorder on and off, and asking the child to talk about specific events that took place in the past and about specific hypothetical events. (This was another attempt to elicit past tense forms, because in English, as in many languages, the inflection called "past tense" is also used to express hypothetical events in certain contexts, as in *I wish I knew the answer* or *If I won the lottery.*) As Kuczaj points out, "the recorded conversations may then not be representative samples of normal conversation."

Kuczaj does not mention whether, in recording Abe, he tended to use the same manipulations that he instructed the parents of his cross-sectional sample to engage in. The transcripts themselves suggest he may have. We tested for a conversational style involving a high degree of past tense elicitations in Abe's transcripts. Using the CLAN program *combo*, we searched for occurrences in Abe's parents' speech of three representative sequences of words that would tend to call for past tense forms in subsequent child utterances: *What did*, *what'd*, and *happened.* Abe's parents used these forms in 1.9% of their utterances, more frequently than 21 of the other 24 sets of adults in our samples, whose mean was 1.1%.[29]

It is unclear whether Kuczaj's selection criteria (or the parent's criteria for when to begin recording) adventitiously correlated with tendencies to overregularize. However, the instructions to parents to elicit past tense forms from their children could have elevated overregularization rates, for reasons we examine in the following section.

Discourse Patterns that Can Elevate Estimates of Overregularization Rates

There are three reasons to think that efforts to elicit past tense forms from children might cause them to overregularize more than they would in their spontaneous speech.

---

[29]Indeed, two of the children whose parents exceeded Abe's in past tense elicitations, Allison and Nat, were considerably younger, and younger children tend to be prompted more often in conversation than older children; see Marcus, 1991).

First, if children have any tendency to avoid forms in their spontaneous speech that they are uncertain of (in particular, the past tense form), then when the parent chooses the verb instead of the child in leading questions, errors may be more likely, for example, in the following dialogue between Abe and his mother (from his 98th sample):

Mother: And what did you choose to do?

Abe: I choosed to make cookies.

Second, by eliciting descriptions of hypothetical and past tense events, the adult may be creating a discourse situation in which the child feels more compelled on pragmatic grounds to mark tense than he would be otherwise. As we discuss in Chapter 12, Abe and Kuczaj's (1977a) cross-sectional sample had unusually high levels of compliance with the requirement of English syntax that past tense be marked in main clauses describing past or hypothetical events (i.e., they seldom used the stem form of verbs in obligatory past tense contexts), and overregularization tends to be accompanied by such high levels of past tense marking. The following bits of dialogue give a flavor of this possibility for past and hypothetical contexts, from Abe's 34th transcript and 144th transcripts, respectively:

Father: Abe do you remember what all you did yesterday?

See if you can tell me what all you did yesterday.

Abe: We played the new games we fixed the wagon and we opened gifts.

Father: What did you do outside?

Abe: Hide. [played hide and seek]

Father: You hide? you hid?

Abe: Uhhuh and I count.

Father: You counted?

Abe: uhhuh.

Mother: Abe was it ... and he counted and came and looked for us.

Father: What happened then?

Abe: I finded Renee.

.....

Father: What would have happened if they couldn't have found any water?

Abe: They gotted a hose.

Third, because of the syntax of English, questions without auxiliaries require insertion of *do* and the verb in stem form. If the child's representation of the stem form is primed by its appearance in a leading question, the stem could become unusually available for the regular inflection process, and relatively less liable to being blocked by the irregular (see Stanners, Neiser, Hernon, & Hall, 1979, and Fowler, Napps, and Feldman, 1985, for evidence that stem forms and regularly inflected forms prime each other).[30] In contrast, spontaneous use of irregular past tense forms might involve activation of the entire lexical entry of a verb and the feature "past tense," directly indexing the stored past tense form, rather than first activating the stem form. The effect might underlie exchanges like this one in Abe's 150th transcript:

Father: What did you hear?

Abe: I heared something like the TV.

There is suggestive evidence that higher levels of past tense elicitation in general may indeed increase overregularization rates. The number of times Abe heard *what did*, *what'd*, and *happened* per month correlated positively with his overregularization rate for that month, $r(31) = = .35$, $p < .05$. (Across children the overregularization rate correlates

[30]Furthermore, providing the stem may have actively suppressed retrieval of the irregular version. Presenting an adult subject with a subset of a category of remembered words can impede retrieval of the rest (Slamecka, 1969). We thank Endel Tulving for pointing this out to us.

positively, though nonsignificantly, with the number of these forms per parental utterance; $r(23) = .18$, $p = .20$.) To see if these correlations might reflect a causal relation, we extracted all of Abe's stem overregularizations and correct past tense forms (excluding *get*, *read*, and the no-changers).[31] The preceding parental utterance was then scored manually either as a clear past tense elicitation (e.g., "Tell me what you did," "What did you [verb]?," and so on) or as some other sentence type. Overregularizations were significantly more likely after past tense elicitations than other kinds of sentences (27% versus 19%; $X^2(1) = 9.64$, $p < .005$).

We also tested for one of the specific mechanisms by which certain kinds of past tense elicitations might increase overregularization: the priming of the child's representation of the stem form by its immediately preceding use in a parental question. Abe was significantly more likely to produce a stem overregularization of a particular verb (compared to the correct irregular form) when his parents used the stem form of that verb in the conversational turn immediately preceding the overregularization than when they did not use the stem form in that turn (32% versus 21%, $X^2(1) = 5.60$, $p < .05$).

To see if these discourse influences operate generally in children, we ran similar tests on the much smaller set of overregularizations from Adam, Eve, and Sarah. We collected all the stem overregularizations in their transcripts, and paired each one with the nearest correct irregular past tense form in the transcripts. All three children showed a greater overregularization rate following past tense elicitations (9% versus 2% for Adam, 44% versus 15% for Eve, 25% versus 18% for Sarah); the difference was significant by a Chi-square test for Eve individually, and in a meta-analysis combining the probabilities of the Chi-square tests for the four children ($p < .01$). However, we were unable to confirm the

---

[31]Only stem overregularizations, not past overregularizations, were extracted, because this analysis was done simultaneously with the one described in the following paragraph, which requires stem overregularizations.

operation of a specific mechanism contributing to this effect that we tested for, namely the priming of the stem by an immediate prior occurrence in the parent's prompt. Tests of this effect were inconclusive: Adam's and Eve's samples contain more overregularizations of a verb after the parent had just used its stem; Sarah's contained fewer.

In sum, there are reasons to believe that the higher overregularization rates among the children Kuczaj studied, including Abe, are partly artifactual. The data are, as Kuczaj's noted, "not representative samples of normal conversation"; rather, they appear to be part way between naturalistic speech and experimentally-elicited speech. Our analyses of discourse effects on overregularization, together with our review of data from actual elicitation experiments in the next section, suggests that Kuczaj's methodology might have led to systematic overestimates of children's spontaneous rate of overregularization.

Note that we are not suggesting that the entire difference between Kuczaj's subjects and those in the rest of the literature is artifactual; in Chapter 12, we will examine the possible effects of high levels of marking tense on overregularization, a difference between Kuczaj's subjects and the other children that is of some theoretical interest. Note as well that these analyses are an attempt to resolve the discrepancy between Kuczaj's data and the others in the literature, and should not be interpreted as a criticism of Kuczaj's methodology in general. Three of the four hypotheses he was testing in Kuczaj (1977a) had nothing to do with overall overregularization rates, but compared the two different kinds of overregularization errors (*eated* and *ated*), and the relative overregularization rates of verbs with different properties. For these purposes there were good reasons to have sought samples with large numbers of past tense forms, just as we focus on Abe in our investigation of the time course of individual verbs in Section 10.[32] It is only when treating Kuczaj's data as estimates of children's overall overregularization rates in spontaneous speech that the nonrepresentativeness of the samples must be taken into account.

Previous Estimates of Overregularization from Elicited Production Experiments

Several experimental studies have elicited past tense forms in sentence completion tasks using existing irregular English verbs. For example, in the experiment by Kuczaj (1978), the overregularization rates for the groups of 3-4 year-olds, 5-6 year-olds, 7-8 year olds were respectively 29%, 49% (42% stem overregularizations, 7% past+*ed* overregularizations), and 1%. Bybee and Slobin (1982a) found that their third-graders (8;6 - 10;1) overregularized between 2% and 55%, depending on the verb subclass. Marchman (1988) found the following overregularization rates for her different age groups: 4's, 32%; 5's, 33%; 6's, 22%; 7's, 10%; and 9's, 5% (calculated from her Table 1, based on 76% of the test items being irregular, as mentioned in her text.) Note that in all such studies the overregularization rates are generally less than 50%; once again there is virtually no evidence that young children overregularize exclusively or in free variation with correct irregular forms.

Of course the overregularization rates in elicited production tasks are still far higher than those obtained from spontaneous speech, but the two kinds of estimates are not

[32]Kuczaj's tests of his fourth hypothesis, however, may have been affected by his methodology, though of course he was not aware of the effects we now point out. Kuczaj disagreed with Brown's claim (1973) that irregular past tense marking was acquired before regular past tense marking, which Brown had based on differences in the relative ages at which the two kinds of marking were supplied in obligatory past tense contexts more than 90% of the time. Kuczaj noted that the finding might depend on whether overregularizations had been counted as incorrect forms or ignored when the percentages were tallied, which had not been made clear in Brown (1973). (In fact, they were ignored; Brown, Cazden, & de Villiers, 1971.) Kuczaj counted the overregularizations twice: as incorrect irregular past tense forms, and as correct regular past tense forms. The resulting percentages of correct use in obligatory contexts showed an advantage to regular marking, contrary to Brown's claim. Thus Kuczaj's "nonreplication" was partly just due to differences in definitions, but also due to how these differences interact with his special recording circumstances. Because these circumstances may have led to unusually high overregularization rates, Kuczaj's estimates of irregular marking were thereby depressed, and his estimates of regular marking inflated.

comparable. Bybee and Slobin (1982a), Stemberger and MacWhinney (1986), and Prasada, Pinker, and Snyder (1990) found that adults, when put under time pressure, are prone to making overregularization errors at even higher rates than children (from 6% to 31% of the time in the Bybee-Slobin study, depending on the subclass), presumably because of a greater likelihood of retrieval failure. It is plausible that many children feel that they are under pressure in experiments even if it is not explicitly stated. Furthermore if children ever fall into a strategy of treating each experimental item as a pure sound, rather than as a word they know, it essentially becomes a novel form and regularization is the most accessible option.

But most important, in all such tasks children are being supplied with the stem itself seconds before they are asked to supply the past form (e.g., "This is a girl who knows how to *swing*. She did the same thing yesterday. She ____"). This contrasts with naturalistic settings in which children produce a past form for an irregular in response to a mental representation of the verb's meaning plus the feature for past tense; the phonetic form of the stem need never be activated. Thus experimental elicitations of irregular past tense forms using the stem as a prompt, like parents' leading questions containing the stem, are likely to prime the child's representations of the stem form and possibly suppress the irregular past, leading to an increased likelihood of overregularization, an effect we were able to document for Abe in the preceding section.

Previous Estimates of Overregularization from Judgment
and Correction Experiments

Another source of data that might be thought to show that children are indifferent to the past tense forms of their irregular verbs comes from Kuczaj's (1978) judgment task. In one experiment children aged 3 to 9 years had to judge whether any member of a group of puppets "said something silly." One puppet produced a sentence with a correct irregular past tense, a second produced an overregularization, and for verbs other than no-changers, a third produced a past+*ed* form. In a second experiment, children of the same ages produced past tenses for irregular verbs supplied in the future tense (discussed in the preceding section), then judged a puppet's version of the verb (always different from the child's version), and then judged a second puppet's version (the third possible kind of past tense form). Finally, children were offered a forced choice among the three versions, and asked which of the three they thought their mother would use. In many conditions, overregularizations were judged as acceptable a large proportion of the time, as high as 89% for stem overregularizations for the youngest children in the first experiment.

However here too the data are not comparable to overregularization rates from spontaneous speech. Grammaticality judgment is a signal detection task, and it is fallacious to assume that every time a child accepts or fails to correct a given form, the child's grammar deems it well-formed. Rather, just as with all yes-no data, the perceived payoffs for hits, misses, false alarms, and correct rejections affect rates of saying "yes." For children in an experimental setting this could involve a variety of demand characteristics such as the perceived politeness of rejecting or correcting another creature's language more than a given proportion of the time. In the language of the Signal Detection Theory, this defines a "criterion" or bias for saying "yes" that is superimposed on their "sensitivity" in internally representing grammatical and ungrammatical utterances as different, which we can assume is a probabilistic process. Lacking direct manipulations of bias, the best one can do in determining whether children have knowledge of irregular pasts is to compare their "yes" rate for correct irregulars versus incorrect overregularizations; if the former are higher, children must be discriminating between them (see Grimshaw and Rosen, 1990). Kuczaj's data provide 15 opportunities to make such comparisons: three age groups in Experiment 1, each of which was asked to judge overregularizations of no-change irregulars, and to judge overregularizations of other irregulars; and three age groups in Experiment 2, each of which was asked to judge stem overregularizations, to judge past overregularizations, and to choose among their favorite from among the three. (The children's choice of their mother's favorite form was almost always identical to their own choice, and thus is not an independent data set.) Of these 15 comparisons, only 1 involved a failure to discriminate irregulars from overregularizations: the middle age group (5-7) in Experiment 2 preferred past overregularizations over irregulars or stem overregularizations in the forced choice task (though they did not even produce many such forms in the elicited production task, as noted above). In other words, the judgment data confirm that children systematically favor irregular forms as the preferred past tense version of irregular verbs (see also Lachter & Bever, 1988).

More recently, Cox (1989) told children that a puppet "was learning to talk but was having trouble with some of his words," and the child was asked to help him say the correct words. Twelve sentences, each with an overregularized noun or verb, were provided. Children were not asked to judge the sentences, and there were no correct irregulars among the experimental stimuli, so we cannot assess children's discrimination abilities from the data. Correction performance was surprisingly poor: none of the 6 sentences with verbs was corrected by more than 16% of the children around the age of 5, and none of the 6 sentences with nouns was corrected by more than 28%, except, inexplicably, *tooths*. Since, as Cox herself notes, the children who failed to correct an overregularization did not necessarily use it themselves, a response bias against correcting the puppet too often is a likelier explanation than an absence of knowledge, especially since she told the children that the puppet was having trouble only with "some" of his words, but presented no sentences that were actually correct.

Finally, the consistent findings that overregularization rates are low and that irregulars are preferred to them helps to explain the otherwise paradoxical phenomenon that children who have been observed to overregularize will vehemently correct their parents when they mimic the children's errors (Slobin, 1978; Bever, 1975; see Lachter and Bever, 1988). Similarly, Ervin and Miller (1963) noted that their subjects often corrected their own overregularizations; we do not know of any reports of children correcting one of their irregular past tense utterances to an overregularization.

Summary. Overregularization percentages in the single digits are characteristic of most children, ages, and commonly used verbs. This suggests that children, during the ages at which they are overregularizing, are not failing to discriminate irregular verbs from regulars, nor are they freely alternating between overregularizations of irregulars and the corresponding correct past tense forms, but show a pervasive strong bias for the correct form. Such a bias is predicted by the hypothesis that children block regularization whenever they retrieve an irregular form from memory (MacWhinney, 1978; Pinker, 1984), but is not predicted by any other hypothesis in the literature. Finally, what has been called "U-shaped development" corresponds to the following two events: Before the first overregularization, there is a measurable extended period where all irregular past tense forms are correct, and overregularization tails off gradually during the school age years.

11:
The Relation of Overregularization to Changes in the Number and Proportion of Regular Verbs in Parents' Speech and Children's Vocabulary

In this chapter we review the operation of the Rumelhart-McClelland model and how it models children's U-shaped developmental sequence. Then we examine the empirical assumptions justifying the sequence of inputs that Rumelhart and McClelland trained the model on. We test whether these assumptions are reasonable using the available data from the spontaneous speech of Adam, Eve, and Sarah, the three children we have been focusing on who displayed a U-shaped sequence in their longitudinal development.

Why the Rumelhart-McClelland Model Begins to Overregularize

The core of the Rumelhart-McClelland model is a pattern associator network that takes a phonological representation of the stem as input and computes a phonological representation of the past tense form as output. The pattern associator consists of two layers of nodes -- a set of input units that are turned on in patterns that represent the sound of the verb stem, and a set of output units that are turned on in patterns that represent the sound of the verb's past tense form -- and weighted connections between every input unit and every output unit. Each unit corresponds to a sequence of phonological features, such as a high vowel between two stop consonants, or a back vowel followed by a nasal consonant at the end of a word. The word itself is represented solely by the set of feature sequences it contains. When a set of input nodes is activated, each node sends its activation level, multiplied by the link weight, to the output nodes it is connected to. Each output node sums its weighted inputs, compares the result to a threshold, and probabilistically turns on if the threshold is exceeded. The output form is the word most compatible with the set of activated output nodes.

During a learning phase, the network compares its own version of the past tense form with the correct version provided by a "teacher," and adjusts the strengths of the connections and the thresholds so as to reduce the difference between the actual state of each output node and the correct state. By this process of recording and superimposing contingencies between bits of sounds of stems (e.g., the distinctive features of endings such as *-op* or *-ing*) and bits of sounds of past tense forms (e.g., the features of *-opped* and *-ang*), the model improves its performance over time, and can generalize to new forms on the basis of their featural overlap with old ones. The model contains nothing corresponding to a word or rule, and thus makes no qualitative distinction between regular and irregular mappings; both are effected by connections between stem sounds and past sounds.

Rumelhart & McClelland's explanation of the sequence of overregularization flowed from the ways in which PDP models generalize. In building such models, there are numerous ways to bias it toward conservative recording of individual input items, toward liberal overgeneralization according to frequent patterns, or some combination. The challenge was to duplicate the child's transition from conservatism to overgeneralization in a single model. Rumelhart and McClelland proposed a simple and ingenious hypothesis. Not only are irregular verbs high in frequency, but the reverse is true as well: the verbs highest in frequency are irregular. For example, the top ten verbs in Kucera & Francis's (1967) frequency list are all irregular. If children acquire verbs in order of decreasing frequency, they will develop a vocabulary with an increasing proportion of regular verbs as they begin to run out of the high-frequency irregulars and encounter more and more regular verbs. In particular, Rumelhart and McClelland assumed that at some point in development the child shows "explosive" vocabulary growth, which would result in a sudden influx of a large number of regular verbs. Because the regular pattern will be exemplified by many different verbs, the learning procedure will strengthen many links between stem features and the features defining the *-ed* ending. The effects of these newly-modified link weights could overwhelm the existing weights on the links between idiosyncratic features of irregular stems and the idiosyncratic features of their pasts, resulting in overregularization. As the irregulars continue to be processed, the discrepancies between the overregularized and teacher-supplied correct forms will be registered, and the crucial idiosyncratic links will be strengthened over time, eventually allowing the irregular forms to reappear.

Given these assumptions, Rumelhart and McClelland were able to model the developmental sequence with one additional assumption: The vocabulary explosion occurs after the child has just acquired his tenth verb. Their ten-verb decision results in two training phases. First, the model is presented with the ten highest-frequency verbs (excluding *do* and *be*, which can also be auxiliaries), of which only 2 (20%) happen to be regular, 10 times apiece. Then the model is presented with that list plus the 410 next-most-frequent verbs, constituting a set in which 80% of the verbs are now regular, 190 times apiece. In Phase 1 the model learned the 10 verbs successfully; when Phase 2 begins on the eleventh cycle, and it is suddenly swamped with regulars, the model overregularizes the irregulars. The recovery process begins immediately, and reaches asymptote shortly before the 200th epoch (see Figures 16 and 17, taken from Rumelhart & McClelland, 1986).

Figures 16 and 17 here

The Rumelhart-McClelland model challenges the traditional account of overregularization, which depended on separate rote and rule mechanisms, in favor of a single mechanism that begins to overregularize because of an influx of newly-acquired regular verbs, a presumed consequence of a vocabulary growth spurt. Let us call this explanation of the cause of overregularization the Vocabulary Balance hypothesis; it is also a feature of the more recent network simulations by Plunkett and Marchman (1990).

Pinker & Prince's Critique

Pinker and Prince (1988; see also Prince and Pinker, 1988) examined Rumelhart and McClelland's assumptions about development. Rumelhart and McClelland cited Brown

(1973) in support of their assumption of a vocabulary spurt near the onset of overregularization, but Brown did not discuss vocabulary acquisition at all. According to standard sources (see, e.g., Ingram, 1989) children's "word spurt" usually occurs at 1;6, about a year too early to account for the onset of overregularization for most children, which occurred at a mean age of 2;5 for the 7 children examined in Section 10. Pinker and Prince examined Brown's (n.d.) vocabulary lists for Adam, Eve, and Sarah, from 5 evenly spaced samples spanning the overregularization sequence, plus a fourth child in the one-word stage. They found no explosive growth in vocabulary near the onset of overregularization, nor, more significantly, an increase in the percentage of the child's vocabulary samples that was regular: The proportion regular stayed around 50% before, during, and after the onset. Pinker and Prince also cited data (partly from Slobin, 1971) suggesting that the proportion of regular verb tokens among all verb tokens in parental speech to children is about 20% to 30% during overregularization, nowhere near the 80% proportion that Rumelhart and McClelland used to override the irregular patterns. They argued that an endogenous transition from rote to rule is still required to account for the data, as in the traditional account.

Pinker and Prince's (1988) data showing that the proportion of regular verb types in children's longitudinal samples' stays at around 50% seems paradoxical at first: if there are only 180 irregular verbs and thousands of regulars, isn't an increase in the percentage of regular verbs a mathematical certainty after the 180th irregular is acquired, and a statistical near-certainty well before that? The answer is that Pinker and Prince's type estimates were from fairly small samples (about 700 utterances per child per stage) and hence were not pure estimates of type frequency, but something combining type and token frequency: Types with higher token frequency were more likely to have been sampled. Because the token frequency of irregulars is much higher than that of most of the regulars, it is possible that when children learn lower frequency regular verbs, they may not displace the earlier acquired irregulars. *Permit*, *understand*, *remember*, *misbehave*, and so on may compete among themselves for air time in children's speech, leaving general-duty verbs like *come*, *go*, *take*, *put*, *eat*, and so on to occupy a constant proportion of verb slots in conversation throughout development.

Because Pinker & Prince's data reflected both type and token frequencies, the force of their critique is uncertain. To evaluate the Vocabulary Balance hypothesis, then, one must first establish whether it is the proportion of regular types, tokens, or some other index that is relevant. Two issues must be addressed: What is the psychological event that corresponds to an episode of network learning, according to Rumelhart and McClelland's theory? And what kind of changes in the schedule of learning episodes cause overregularization in pattern associator networks?

What is a learning episode?

Rumelhart and McClelland make the following assumptions about the real world events that correspond to a learning episode:

> The [simulation] run was intended to capture approximately the experience with past tenses of a young child picking up English from everyday conversation. Our conception of the nature of this experience is simply that the child learns first about the present and past tenses of the highest frequency verbs; later on, learning occurs for a much larger ensemble of verbs, including a much larger proportion of regular forms.
>
> Although the child would be hearing present and past tenses of all kinds of verbs throughout development, we assume that he or she is only able to learn past tenses for verbs already mastered fairly well in the present tense. This is because the real learning environment does not, in fact, present the child with present-tense/past-tense pairs. Rather, it presents the child with past-tense words in sentences occurring in real-world context. The child would therefore have to

generate the appropriate present tense form internally with the aid of the entire sentence and context, and this, we suppose, requires that the child already know the present tense of the word. (Rumelhart and McClelland, 1987, p.222).

The assumption here is that an episode of learning consists of hearing a past tense form, using the context to recover its corresponding stem from the mental lexicon,[33] feeding the stem into the internal pattern associator, comparing the output with the past tense form actually heard, and adjusting the weights in response to discrepancies. A stem-past pair would be fed into the model only when an adult used the past *and* the child possessed the stem in his or her vocabulary (and knew it was related to the past form).

This means that the proportion of regulars fed into the past tense learning system would be determined by the proportion of occasions that the parent used a regular past tense that the child already possessed in stem form. However, it is impossible to tell from transcripts exactly when this conjunction of parent's use and child's knowledge occurs. Instead, there are three ways to estimate the relevant proportions indirectly, each with different assumptions.

For all three, it is useful to assume that all verbs have an approximately constant distribution of uses in different tenses, so we can collapse across tenses and increase sample sizes, reducing the danger of underestimating the number of regular verbs. Counting all verbs, not just past tense forms, also reduces the size of a possible confound: if one were to find a positive correlation between the increase in the number of regular past tense forms and the tendency to overregularize, it could reflect the effects of a newly-acquired regularization process on the ease of generating past tense forms of regular verbs, rather than the effects of having many regular verbs on the development of an ability to

[33]Rumelhart and McClelland refer to the input form as the "present," but the present tense form would include an irrelevant *-s* affix for the third person singular; "stem" is actually what they had in mind.

regularize.[34] In fact, we have found that irregular verbs take up a somewhat larger proportion of past tense tokens than of total verb tokens (about 85% versus 65-75%), but this difference only strengthens the conclusions we will be making on the basis of all verb tokens.

The first method assumes that children produce regular and irregular verbs in approximately the same proportions that they process regular and irregular past tense forms in their parents' speech (that is, it assumes that if a child uses a verb, he or she knows it, and that children in conversation with parents will use different verbs in roughly the same proportion as their parents). Under this assumption, *the proportion of verb tokens that are regular in the child's speech* indirectly estimates the proportion of regular learning episodes.

Second, since the occurrence of a parental token is necessary for a learning episode to take place, if we assume that children know a constant proportion of the verb tokens their parents address to them we can measure *the proportion of verb tokens that are regular verbs in the parent's speech.*

In practice, Rumelhart and McClelland ignored token frequency entirely in assembling the training set for their model: every verb was fed in the same number of times, once per epoch. This assumes that a third measure, the proportion of regular verb types among all verb types, is the relevant factor -- though it is inconsistent with their psychological interpretation of a learning episode, which would be driven by parental

---

[34]Lachter and Bever (1988) point out that the direction of a cause and effect relation between possessing regular verbs and overregularization cannot completely be resolved, even if one were to examine possession of regular verbs per se rather than mastery of the past tense form of such verbs. The problem, they note, is that once a child has developed the ability to generate and analyze regular past tense forms productively, he or she no longer has to memorize such forms from parental speech. As a result, the learning events and memory space that beforehand would have been dedicated to acquiring regular past tense forms can now be applied to the learning of brand new verbs, and such verbs should be acquired at a more rapid rate. We will not pursue this possibility further, though it would be an important consideration if one were to find the relevant correlation.

tokens. Rather, the teaching schedule they actually modeled is more consistent with some kind of off-line learning, fed by a preprocessor: The child takes a pass through his entire verb lexicon, feeding each stem-past pair into the pattern associator once per scan. If we entertain this interpretation of a learning episode, which corresponds literally to Rumelhart and McClelland's learning schedule, rather than the token-driven interpretation they discuss, we can test the Vocabulary Balance hypothesis by trying to estimate the *the proportion of verb types in the child's vocabulary that are regular*. Note that it is not valid to use the percentage of a child's vocabulary that is regular as a surrogate for the number of on-line learning episodes even if it turns out that the proportion of regular verbs among parental tokens is constant. That is, one cannot assume that the proportion of regular learning episodes is determined by the proportion of the child's vocabulary that is regular, because this larger regular vocabulary could correspond to a larger number of regular types that the parent is cycling through a constant number of regular tokens in his or her speech, leaving the proportion of regular learning episodes constant.

How do changes in learning episodes lead to overregularization?

Assuming we know what a learning episode is, what kind of changes in the distribution of learning episodes lead to overregularization, according to the Vocabulary Balance hypothesis?

Type versus Token Frequency

First, it is clear that both type and token frequency have important consequences. The Rumelhart-McClelland model overregularized because it changed its connection strengths with each input pair in a direction that reduced the discrepancy between computed and input past forms. After the first epoch in which the model was suddenly bombarded with regulars, about 80% of the changes the model made were designed to make it more likely to generate regular forms, because 80% of the inputs were new regulars. Many of the changed connections involved links from phonological features that were also shared with irregulars (since most irregulars are phonologically similar in some way to at least some regulars). Because the network did not have enough specific feature units to register each verb on its own set of units, the overlap was high enough that each irregular was represented by many units whose links had just been adjusted to help produce the regular ending, and overregularization resulted. This effect would obviously depend strongly on the number of regular types, because the wider the range of regular forms that are fed in, the greater the probability that a given phonological feature of an irregular verb will be shared by some regular verb and hence develop stronger links to the incorrect regular pattern. But this effect can also be mitigated by token frequency: if, say, each irregular had been repeated 4 times for each regular (reflecting the real-world higher token frequency of irregulars), the links that joined features unique to the irregulars to their corresponding irregular past forms would have been strengthened several times during the epoch to reduce the errors with such forms, and overregularization would be less likely. Indeed, one of the noteworthy properties of the Rumelhart-McClelland model is its distributed phonological representation of words, with no units dedicated to words per se (see Pinker & Prince, 1988), so there is no physical basis for a distinction between types and tokens in the model at all. Only feature-to-feature mappings, whether they be from a single word or a set of similar words, are represented. The actual behavior of the model will depend on the number of regular types, the phonological range of the regular types, their degree of overlap with irregulars, the token frequencies of both irregulars and regulars, and other factors. In any case it is clear that overregularization of an irregular does depend on the ratio of regular to irregular tokens, and hence is relevant to testing the Vocabulary Balance hypothesis.

Proportion versus Changes in Proportion

Second, the percentage of regular learning episodes at a given time is not the relevant factor in predicting overregularization. In the Rumelhart-McClelland model, unlike in children, the process of recovery from overregularization begins immediately after its onset (see Figure 17), correct irregulars predominate within a few epochs, and at asymptote they are produced most of the time, all with a constant level of 80% regular learning episodes. This is an obvious property of any model that is designed to perform correctly at asymptote: Even with the most unfavorable proportion of regular episodes, the irregulars must eventually reassert themselves. Overregularization is a short-term consequence of the *increase* in the percentage of regular episodes with development. Though a properly-designed model could learn to overcome any particular level of dominance of regulars, this adjustment cannot take place instantaneously, and influxes of regulars will cause temporary overregularization, before the crucial links between nodes unique to an irregular and its idiosyncratic past have been sufficiently strengthened. For this reason, the difference between the rapid recovery of the Rumelhart-McClelland model and the protracted period of overregularization of children does not speak against the model. It is possible that as children learn more and more words, new regulars are constantly washing over them; no sooner do they adjust their irregulars to the leveling effect of one wave of regulars than a new wave comes in. Thus the proper test of the Regular Vocabulary Balance hypothesis involves a correlation between the most recent *increase* in regular learning episodes and the current rate of overregularization; this is why Rumelhart and McClelland appeal to a period of explosive *growth* in vocabulary to trigger overregularization.[35]

Proportion of Regular Verbs versus Number of Regular Verbs

Third, for the analyses in which types are being examined, it may not be the proportion of verb types that are regular that is the relevant predictor. The problem is that the competition in pattern associators is not between the regular pattern and a single irregular pattern shared among all the irregulars. Rather, irregulars are different from each

---

[35]In recent experiments on the behavior of connectionist networks at learning inflectional mappings in sets of artificial verbs with different training schedules and vocabulary mixtures, Plunkett and Marchman (1990, 1991) have confirmed that both token frequencies and rate of vocabulary increase have direct effects on the tendency of standard connectionist models to produce outputs analogous to overregularizations.

other, not just from the regulars. Imagine that at one stage, there are six different irregulars, each with a different change (e.g., *go-went*, *come-came*, *hit-hit*, etc.) and six regulars. At the next stage, there are twelve different irregulars, each with a different change, and twelve regulars. The proportion of regulars in the sample remains the same, but the ratio of regulars to *any particular irregular pattern* has doubled. Therefore in this scenario it is the number of regulars, not the proportion of regulars, that would predict overregularization.

The preceding scenario is not fully accurate, however: the ratio of learning episodes for the regular pattern to a given irregular pattern would be identical to the number of regulars only if every irregular were *totally* idiosyncratic. But virtually all the irregulars share their patterns of change with other irregulars, so the calculation is too extreme. Consider a scenario that is extreme in the other direction: the six irregulars in phase one fall into three classes (e.g., *sing-sang*, *ring-rang*, *feed-fed*, *breed-bred*, *wear-wore*, *tear-tore*), and the new irregulars in phase two fall into the same classes (e.g., *spring-sprang*, *lead-led*, *swear-swore*). Here the ratio of regulars to any irregular vowel change pattern is 3:1 in both phases, and we would expect overregularization to be less likely; only a change in the proportion of regulars would clearly induce it.

In reality the situation is likely to be somewhere between these extremes because while English irregulars do fall into a restricted number of kinds of change, we would expect the number of patterns in a child's vocabulary, not just the number of irregulars per pattern, to increase somewhat with development. Therefore it is not clear whether overregularization rates should be correlated with the proportion of total types that are regular (appropriate if all new irregulars fall into old patterns and hence protect old irregulars), or the number of types that are regular (appropriate if each new irregular is unique), and we will examine both.

Summary of Tests

We will examine whether there are increases over time in the proportion of verbs that are regular verbs among the child's tokens, the parents' tokens, and the child's types, and if so, whether such increases are related to the child's tendency to overregularize. The relations will be tested at two levels of temporal detail.

First, we will compare the monthly rate of increase in each vocabulary factor for the months before the first overregularization with the monthly rate of increase for the months during which overregularization is taking place (i.e., the first month containing an overregularization and all the months after it.) This is an objectively specifiable dividing line but requires some further comments. For Eve, there are a small number of tokens before the first overregularization, and recall that there is no statistical evidence that she in fact underwent a transition between stages (see Section 10). For Sarah, the dividing line could be questioned. Brown (n.d., see Pinker & Prince, 1988) noted that Sarah produced a past tense of *hear* at 2;10 that was literally pronounced as an overregularization (*heared*, with a schwa instead of the *r*), but he worried whether it could have been an odd pronunciation of *heard*. The first completely unambiguous overregularization does not occur in her transcripts for another 5 months. However, the hypothetical pattern of distortion that Brown considered is not independently motivated by Sarah's other mispronunciations: for example, at 3;2 she pronounced *hurted* as *hahted*, suggesting that if *heard* was intended but mispronounced it would have surfaced as *hahd*, not *hea-id* (Alan Prince, personal communication). Furthermore, excluding this datum has the effect of weakening the evidence for the Vocabulary Balance hypothesis. Therefore we will not second-guess the transcription, and will count *heared* as Sarah's first overregularization. In any case, the first recorded overregularizations of all three children should not be taken as the literal moment of arrival of the ability to overregularize, but as a partly-arbitrary dichotomization of the developmental span into a period when overregularizations are likely to occur and a preceding span when overregularizations are unlikely to occur.

A more precise test comes from correlating the monthly rate of increase in a vocabulary measure between month t and month $t + 1$ with the overregularization rate at time $t + 1$. This analysis combines the factors that differentiate the pre-overregularization stage from the overregularization stage and the factors that cause overregularization to be more frequent in one month than another during the overregularization stage. According to the Rumelhart-McClelland model, there is no qualitative difference between the two. Because longitudinal transcripts from a given child are not independent sampling units, the correlation coefficients are best treated as descriptive statistics; the meaning of tests of significance is not clear. Fortunately, the issue will seldom arise, as the sign of the correlation coefficients we find will usually obviate the need for significance testing.

Parental Tokens

The proportion of regular verb tokens among all adult verb tokens is plotted for Adam, Eve, and Sarah in Figures 18, 19, and 20. As mentioned in the Method chapter, the figures include not only the child's parents but the other adults speaking to the child in the transcripts. For all three children, a bit more than a quarter of the parental verb tokens were regular, and this did not change over the course of development. As mentioned, the reason the proportion is constant is that most of the high-frequency verbs that are indispensable for casual conversation are irregular, and do not move aside to make way for the more numerous but lower-frequency regulars.

The proportions are similar before and after overregularization begins: For Adam, 30% before, 29% during; for Eve, 29% before, 25% during; for Sarah, 25% before, 26% during. More importantly, the rate of change in the proportions was not systematically larger during the overregularization period: For Adam, +1.2 percentage point per month before, -0.1 during; for Eve, +0.9 versus -0.7; for Sarah, +0.2 before and after. The correlations between the rate of monthly change in proportion regular and the child's overregularization rate are in the wrong direction for Adam (-.40) and Eve (-.29) and close to zero (.03) for Sarah.

Figures 18-20 here

Child's Tokens

The proportion of the child's tokens that consists of regular verbs is plotted on the same axes as their adults' proportions, also in Figures 18-20. The proportion oscillates between 26% and 45% for Adam, actually declines for Eve from around as high as 54% (possibly sampling error) to a steady state between 23% and 30%, and shows some early dips to the teens for Sarah before mainly oscillating within the 20% - 40% range. The reasons for regular verbs being used a steady minority of the time are no doubt the same as for the adults.

The proportion of tokens that are regular does not systematically increase after overregularization begins: 37% versus 34% for Adam; 50% versus 27% for Eve; 25% versus 30% for Sarah. The monthly rates of change for these figures are actually higher before overregularization than during: +2.4 percentage points versus -0.3 percentage points for Adam; +7.9 versus -3.8 for Eve; +4.2 versus -0.6 for Sarah. The correlations between monthly change in percent tokens regular and the child's overregularization rate are -.03 for Adam, -.22 for Eve, -.10 for Sarah.

Child's Types

As mentioned, measures based on the child's types are the least psychologically realistic measures to focus on, because they are only meaningful if rule learning is an off-line pass through the child's vocabulary, which does not correspond to Rumelhart and McClelland's psychological assumptions. But the measures are worth examining both because they represent the form of the Vocabulary Balance hypothesis that is most likely to be consistent with some developmental trend (since the proportion of regular types *must* increase with development) and because it literally corresponds to the training sequence given to the Rumelhart-McClelland model.

Measuring the proportion of children's vocabulary that consists of regular verbs is an extremely difficult problem, for it faces the notorious pitfalls involved in estimating children's vocabulary size in general (see, e.g., Seashore & Eckerson, 1940; Templin, 1957; Lorge & Chall, 1962; Miller, 1977; Moe, et al., 1982, for extensive discussion.) The source of the problem here is that we are confined to the actual words that children used in samples. Obviously the child will only use a small fraction of his or her total vocabulary in any given sample. Since high-frequency verbs are more likely to appear than low-frequency verbs, the number of low-frequency verbs will be systematically underestimated. And since there are more low-frequency regulars than low-frequency irregulars, counting types per sample will systematically underestimate the proportion of regulars (this was a problem with Pinker and Prince's 1988 estimates). There is no completely adequate solution to the problem of measuring children's vocabulary, but there are various estimates that can be examined, and at the very least the direction of changes in the proportion that is regular can be compared for months associated with different levels of overregularization, and the resulting conclusions can be compared across different methods in an attempt to arrive at converging conclusions.

Method 1: Cumulative Vocabulary

One measure that is designed to be generous to low-frequency forms is the child's *cumulative* vocabulary totals. That is, one assumes that the child never forgets. If a word is used in a given month, it is credited to the child's vocabulary from then on. Figures 21-26 shows the children's cumulative vocabulary growth for regular and irregular verbs, and the proportion of cumulative vocabulary that is regular; overregularization rates are included in these graphs for ease of comparison of their developmental courses. Table 5 shows the rate of change of the number and proportion of regular verbs in the vocabularies of Adam, Eve, and Sarah before and during the period marked by the first overregularization; the static vocabulary figures for the month before the first overregularization and the last month of the transcripts are also provided as reference points. By mathematical necessity each child possesses a larger cumulative regular vocabulary later in development than earlier, and as expected, the regulars take up a larger proportion of the child's total verb vocabulary later. However, as the decelerating vocabulary curves (most visible for Adam) suggest, for both the number of regular verbs and for the proportion of verb vocabulary that is regular the rates of increase are much larger for the stages before than during overregularization, contrary to the Vocabulary Balance hypothesis. Similarly, the monthly changes in the number of regular verbs, and in the proportion of regular verbs among all verbs, correlate negatively with overregularization rate for all three children.

Figures 21-26 and Table 5 here

Unfortunately the decelerating vocabulary curve and concomitant negative correlation with overregularization may be a direct consequence, even an artifact, of the cumulative measure for vocabulary. Cumulative vocabulary is equivalent to sampling without replacement. Imagine that parents use 500 regular verbs, with equal token frequencies, when speaking to their children, at all ages. Imagine that every month children attend to and acquire 10% of the verbs they hear, and produce every word they have acquired at least once. At the end of Month 1 their cumulative vocabulary is 50 words. At the end of Month 2 it is not 100 words but only 95 -- the 50 they learned in Month 1, plus 10% of the 450 words that they had not previously acquired in Month 1 (i.e., 45 words; the other 5 they attended to don't count because they had already been acquired.) In Month 3 their cumulative vocabulary will be 136, reflecting the addition of only 40 new words (10% of the 405 remaining), and in Month 4, 172 (36 new words). In other words, new words will be acquired at a faster clip early in development than later, even with a constant learning rate and constant number of words in the environment. If there is indeed a relatively limited set of regular verbs for the child to acquire during the preschool years, this is one reason why the Vocabulary Balance hypothesis could be false.

However these sampling considerations could also mean that the decelerating growth curve is an artifact. Imagine that the child possesses a constant 500 words throughout development, but only manages to produce 50 during a month's worth of samples. Following the same arithmetic as described above, the fact that cumulative vocabulary is a form of sampling without replacement means that we as investigators will spuriously credit the child with having "acquired" fewer and fewer new words with each succeeding month. It is very difficult to tell to what extent the curves in Figures 21-25 represent a genuine sampling effect in vocabulary learning, a sampling artifact in measuring vocabulary from production data, or both. Thus it is important to supplement these direct estimates with some indirect measure that is free of this possible bias.

Method 2: Jackknife Estimates Using "Mark-Recapture" Patterns

There is a family of techniques commonly used in biology and demography for estimating populations sizes from multiple samples. The simplest version is commonly known as *mark-recapture* (see Seber, 1986). Here is an idealized example of how one might estimate the number of squirrels in a forest. Trap 50 squirrels, paint their tails orange, release them, allow enough time for them to diffuse through the forest, trap 50 squirrels again, and see how many have orange tails. If there are 10 such recaptures, then the first trapping session must have represented 10/50 or 1/5 of the total population. Since 50 were trapped initially, the forest population must be 250.

This logic can be applied to vocabulary estimation as follows. A verb is a squirrel, a transcript is a trapping session, and a verb that appears in two successive transcripts has been recaptured. The proportion of verbs at t_2 that also appeared at t_1, multiplied by the number of verb types recorded at t_1, is an estimate of the vocabulary size. Note that this procedure avoids the possibly artifactual deceleration in vocabulary acquisition inherent in cumulative measures. If a child had a static vocabulary of 500 words, 50 of which were recorded in each sample, then the second sample would consist of 5 of the words that appeared in the first sample (1/10 x 50) and 45 new words (1/10 x 450), and the recapture rate of 1/10 (5/50), multiplied by the first sample size (50), would yield the correct figure of 500, and this would be true of every pair of successive samples. Note as well that if new verbs are acquired between t_1 and t_2, the estimate will be an unbiased estimate of the vocabulary at t_2. Imagine an idealized case in which 100 baby squirrels were born between capture and recapture. Orange-tailed squirrels are now recaptured with probability 50/350, so the recapture proportion is 1/7. Seven times the capture sample size of 50 is 350, the true population size at recapture.

Unfortunately, unequal token frequencies for different verbs lead to systematic underestimates if one were to use this procedure unmodified. Imagine that some squirrels are more trap-shy than others -- in a simple case, 40% of the squirrels might be "shy," where "shyness" means that the probability of blundering into a trap (1/5, in our example) is cut in half. One expects the first sample to capture 30 of the 150 bold squirrels, but only 10 of the 100 shy squirrels. Sixteen percent (40/250) of the squirrels in the forest now have orange tails, but the second trapping session will only recapture 7 of them: 6 bold ones (1/5 x 30) and 1 shy one (1/10 x 10). Since the second sample consists of 40 squirrels in all ((1/5 x 150) + (1/10 x 100)), the 7/40 recapture rate, multiplied by the 40 squirrels in the first sample, yields an estimate of 229, 21 less than the true figure. For children's vocabulary, verbs with lower token frequencies are "trap-shy," and because many of them are regular, one would obtain systematic underestimates of total verb vocabulary, regular verb vocabulary, the percent of vocabulary that is regular, and the rate of increase in regular vocabulary (though comparisons of higher- versus lower-growth months might still be roughly accurate).

Biostatisticians have dealt with the trap-shyness problem by applying the "generalized Jackknife" estimator to the mark-recapture methodology, first developed by Burnham and Overton (1978, 1979), and extensively investigated by Otis, Burnham, White, and Anderson (1978). Instead of two trapping sessions, there are *k* of them. The numbers of squirrels that have been captured only once, twice, three times, and so on, are tallied. The effects of unequal capturability can thus be estimated by taking into account the distribution of multiple recaptures. While the simple capture-recapture estimate assumes a uniform distribution of capturability, with the extra information of number of recaptures one can assume that individual capture probability is a random variable from an arbitrary distribution. Otis, et al. (1978) found empirically that the Jackknife estimator produces accurate estimates if many individuals are caught a relatively large number of times - that is, if the multiple recapture rate across samples is high.

In our case this procedure consists of comparing sets of five consecutive transcripts, and counting how many verbs were used in each of the 5 transcripts, how many in only 4, and so on. (To make the estimates for the three children comparable, for Sarah we use 5 consecutive *pairs* of transcripts because her speech was sampled for an hour once a week whereas Adam's and Eve's was sampled for two hours once every two weeks.) As required, many verbs did appear in multiple transcripts from such sets. This set of numbers is fed into the Jackknife algorithm, providing an estimate of vocabulary size for that period. With non-overlapping sets of 5 consecutive transcripts, we obtain independent estimates for different ages. The deceleration in cumulative estimates is eliminated, as is the underestimation inherent in simple mark-recapture estimates.

The estimates span periods of 2 1/2 months rather than a single month, which has both disadvantages and advantages. The growth estimates are temporally coarser, and there is no vocabulary estimate at all for Eve that corresponds exclusively to the period before her first overregularization. However a larger temporal window may catch effects of vocabulary growth that act over longer time spans than the one-month window used so far.

The estimator is not free of complications. Because the kind of context in which the recording takes place is similar in all recording sessions, those verbs most appropriate to those contexts will be recorded more often. (In the ecological literature, it has also been noted that achieving equal capturability is impossible, even with randomized capture locations on each sampling occasion; see Chao, 1987.) Furthermore there is a free parameter that must be decided upon in calculating the estimates: the "order" of the estimate, corresponding to the maximum number of recaptures (out of 5, in our case) that are counted in the calculations. Higher order estimates have lower bias but higher variance; there is a complex procedure for selecting the optimal order for a given estimate. For simplicity's sake, we will uniformly report estimates of order 4. We have found these generally to be the highest of the estimates of different order, especially for regular verbs, and hence they are fairest to the Vocabulary Balance hypothesis. But in any case we also found that the growth curves for different order estimates are almost perfectly parallel, so the correlations we calculate are not notably affected by this choice. With these considerations in mind we can cautiously compare periods of high regular vocabulary growth with periods of low regular vocabulary growth, even if the magnitudes of particular increases and totals are not to be taken as perfectly accurate estimates.

Figures 27-32 and Table 6 here

Figures 27-32 show that the vocabulary estimates obtained from this method are higher for young children and somewhat lower for older children than the cumulative estimates, eliminating the severe deceleration that was inherently unfavorable to the Regular Vocabulary Balance hypothesis. They also do not display the unlikely constant 50/50 regular-irregular ratio that would correspond to interpreting Pinker & Prince's figures as estimates of types.

Overregularization rates are included in the graphs for ease of comparison of their developmental courses. Table 6 shows the rate of growth of estimated vocabulary per interval during the period preceding the first overregularization and the period beginning with it; numbers and proportions of regular verbs at the end of these two intervals are shown as well. The number of irregular types shows a very small increase with time; the number of regulars shows a larger one. As Table 6 shows, for Adam, the rate of increase in the proportion of verbs that are regular is larger during the overregularization stage than before it (1.1 versus -0.2 percentage points per 5-sample interval). This, however, is the only comparison from among all those we have performed that is in a direction consistent with the Vocabulary Balance hypothesis. For Adam the rate of increase in the proportion of verbs that are regular from one interval to the next shows no correlation with the overregularization rate at the end of the interval (r =.004); there was virtually no difference in the number of regular verbs acquired per interval during the overregularization stage than before it (9.7 versus 9.1 new regular verbs per interval); and the sample correlation between the size of the increase between intervals and the overregularization rate for the second interval is negative (r = -.08). Sarah acquired regular verbs at a faster rate before her overregularization period than during it (42.2 versus 10.2 additional verbs per interval), and these increases correlated negatively with their ensuing overregularization rates (r = -.26). Similarly, the proportion of her vocabulary that was regular increased much more rapidly before than during her overregularization stage (8.3 versus 0.4 percentage points per interval), and correlated negatively (-.17) with overregularization rate. For Eve, the temporal coarseness of the Jackknife estimates prevents a before-and-after comparison of vocabulary composition, but both the increase in number of regular verbs and the increase in proportion of verb vocabulary that is regular correlate negatively with overregularization rate at the end of the relevant interval.

Comparison to Previous Estimates of Vocabulary Size and Composition

With all the hazards of vocabulary estimates it would be reassuring to compare ours with previous estimates in the literature that report type and token counts from much larger samples of children's speech.

E. Horn (1925) examined the vocabularies of 80 children from 1 to 6, plus 270,000 tokens from kindergarten and first graders. The combined lists yielded 5,000 words, from which he selected 1,084 that recurred a certain minimum number of times across samples. Of these, 233 are unambiguously verbs, 81 irregular 152 regular, or 65% of types regular. M. D. Horn (1927) amassed 489,555 tokens of the speech of kindergarten children, comprising 7,097 types. She reports the 1003 most frequently used words. The list contains 97 regular verbs and 74 irregular verbs, or 57% of verb types regular. These figures are somewhat smaller than the cumulative and Jackknife proportions we find for Adam and Sarah at the end of their transcripts (which range from 69% to 74% regular), but that is to be expected because both Horns excluded lower-frequency words, which in turn systematically underestimates the number of regular verbs. (When we look at single samples of Adam's and Sarah's speech near the end of their transcripts, which introduces a similar bias, we ourselves get lower figures, near 58%.) However, M. D. Horn provided data for the proportion of verb *tokens* that are regular, which should be unaffected by this sampling bias. The token figures are similar to our estimates: 24,581 regular tokens and 80,370 irregular tokens, or 23% regular.

Even better estimates come from Moe et al.'s (1982) figures for first grade children (mean age 6;9), because they report all verb tokens, not just the most frequent ones. Their lists include 418 regular verb types and 108 irregular types, corresponding to 79% of verb types being regular, and 33% of verb tokens. These figures are comparable with the estimates shown in Tables 5 and 6 for Adam and Sarah at the end of their transcripts: Adam (5;2): 74% cumulative types regular, 71% Jackknife-estimated types regular, 33% tokens regular; Sarah (5;1): 73% cumulative types regular; 69% Jackknife-estimated types regular; 20% tokens regular).

Comparison to the Analyses of Marchman and Bates (1991)

Before we summarize the conclusions of this chapter, we briefly examine contrary claims made in a recent paper by Marchman and Bates (1991). They analyze data on children's vocabulary size and overregularizations, and conclude, "verb vocabulary size is highly predictive of ... the onset of overregularization errors (p. 7)." In particular, they interpret their data as support for a "critical mass hypothesis" according to which the "vocabulary size is related to ... the subsequent onset of overregularization errors when verb vocabularies become sufficiently large" (p. 6). They take this as evidence supporting the connectionist models of Plunkett and Marchman (1990, 1991), which also, they suggest, show critical mass effects. In this section we attempt to resolve the discrepancy between these conclusions and ours.

Marchman and Bates used different methods of vocabulary estimation from ours, different measures of overregularization, and different predictions about the relevant correlations. We examine each in turn before discussing their actual findings.

Method of estimation of vocabulary size and overregularizations. The vocabulary estimation technique employed a checklist given to 1,130 mothers of children between the age of 1;4 and 2;6. Each mother filled out the checklist once, so there are no longitudinal data. One section of the checklist contained 680 word stems, including 46 irregular verb stems and 57 regular verb stems. The other consisted of 21 irregular past tense forms and 30 overregularized versions of these verbs (stem overregularizations of all verbs, plus past overregularizations of some of them).[36]

As Marchman and Bates point out, it is desirable to use a variety of methods of vocabulary acquisition, given the limitations inherent to any single technique such as the ones we have noted and dealt with in this chapter. The main advantage they see for parental checklists is that they provide "more accurate estimates of vocabulary size," presumably

---

[36]These figures correspond to the number of verbs actually listed in their Appendix; their text mentions 20 irregular past tense forms and 31 overregularized forms.

owing to the potentially larger corpus of speech potentially drawn from (namely, however much of a child's speech a parent can Remember). Unfortunately, this putative advantage is vitiated by the fact that Marchman and Bates only provided a small subset of English irregular verbs on their list. For example, Adam at 2;6 and Eve at 2;3 already had produced within their transcripts (a small subset of their total speech) more regular verbs, and more irregular verbs, than Marchman and Bates even included on their checklists (see Figures 21 and 23), and this was true for Sarah by the time of her first overregularization (see Figure 25). Thus the verbs that Marchman and Bates's parents had the opportunity to check off constitute a subset of the verbs that we know (from our transcript data) that children of the relevant ages use. Therefore Marchman and Bates's vocabulary figures, contrary to their claims, are not more accurate than our transcript counts; we know they must suffer from a greater degree of underestimation.

A more substantive objection, however, is that the checklist data pertain exclusively to types -- Marchman and Bates estimate how many verb types were overregularized at least once, but have no data on overregularization rate either within or across verbs. As we shall see when we return to their empirical claims, this is a fatal limitation, confounding their main conclusion.

The logic of the prediction. Marchman and Bates, citing the simulations of Plunkett and Marchman (1990, 1991) predict that overregularization should be triggered by the attainment of some absolute size of verb vocabulary. As a result, they tested the correlation between vocabulary size and the number of verbs overregularized. This contrasts with our tests, which test for correlations between the recent *change* in vocabulary size (i.e., rate of acquisition of new vocabulary) with overregularization.

Examination of the Plunkett and Marchman papers, however, raises questions about the basis for Marchman and Bates's prediction. First, Plunkett and Marchman repeatedly conclude that the *rate* of vocabulary expansion (the factor that we, but not Marchman and

Bates, correlated with overregularizations) is a critical variable in their models' tendency to overregularize (e.g., pp. 12-13, 18). Indeed, they note that their model can attain 100% correct performance on any *absolute* level of vocabulary size (p. 11). Second, when Plunkett and Marchman (1990) do discuss a putative effect of a critical mass in vocabulary, the effect in question is "recovery from erroneous performance" (p. 11), exactly the opposite of the *decrement* for irregular verbs that Marchman and Bates are trying to explain. (Later, they discuss generalization to new stems, but this is still different from overregularization of previously-learned ones). Third, Plunkett and Marchman's claim for a critical mass effect in their model (that is, recovery after vocabulary size reaches 50, constant across different rates of vocabulary increase) bears no obvious relation to the simulation data they present. They cite no numbers, instead inviting the reader to inspect their Figures 1 and 2, but neither figure shows any discontinuity in performance for irregular verbs at the point corresponding to the 50th verb acquired; moreover, their Figure 1 shows clear differences in the recovery profile depending on the rate of vocabulary increase.

The empirical claims. Marchman and Bates report two relevant empirical findings. The first is that overregularizations are rarely reported for children with reported verb vocabularies less than 15-30. They take this as evidence for a critical mass effect, but the reasoning is unsound. If children's vocabulary increases with age and they begin to overregularize at a certain point after vocabulary acquisition begins, by logical necessity any finite data set will yield numbers corresponding to the average and minimum sizes of their vocabulary at the time they begin overregularizing. In particular, the minimum will simply correspond to the vocabulary size of the most precocious overregularizer in the sample. This does not constitute evidence for a critical mass effect, that is, that across variation in other factors such as rate and composition of vocabulary, a given child will begin to overregularize when he or she attains a specific vocabulary size.

The second finding is that vocabulary size correlates well with overregularizations, holding linear effects of chronological age constant in a multiple regression. But this finding, too, is close to being a statistical necessity, not an empirical discovery. Recall that Marchman and Bates's data on overregularization consists of the number of irregular types that are overregularized at least once; they have no data on the overregularization *rate* (the probability that the child will use an overregularization as compared to the correct irregular past). Clearly, a child cannot overregularize a verb that is not in his or her vocabulary. If the child has a vocabulary of 5 irregular verbs, he or she cannot be recorded as having overregularized 6 verbs. If the child has a vocabulary of 50 irregular verbs, he or she can, in principle, overregularize 50 verbs but no more, and so on. Thus even if children overregularize at the same rate regardless of vocabulary size, the number of overregularized types recorded will be almost certainly larger for children with larger vocabulary sizes simply because larger vocabularies yield more logical opportunities to overregularize types. Thus, Marchman and Bates's correlation is an artifact that provides no support for the claim that larger vocabularies predict a greater *likelihood* of overregularizing.

In sum, while there may be promise in the use of parental checklists as a corroborating method of estimating vocabulary size, they have no a priori advantages over direct counts from spontaneous speech transcripts, and as Marchman and Bates have deployed them, they have a number of disadvantages. Moreover, Marchman and Bates's conclusion -- that children, like connectionist models, overregularize because they attain a critical vocabulary size -- must be rejected. First, connectionist models have not been shown to overregularize because they attain a critical vocabulary size; rather, their overregularization is mainly caused by high rates of input of regular verbs, which is the motivation for our own analyses discussed in this chapter. Second, children have not been shown to overregularize because they attain a critical vocabulary size; the correlations that Marchman and Bates report are artifacts.

Chapter summary. Our estimates of children's types, adults' types, and children's tokens provide virtually no support for the hypothesis that overregularization is triggered by increases in the number or proportion of regular verbs available to the child. Regular verbs remain a roughly constant proportion of adults' and children's conversational tokens, and never dominate. Regular types -- which in any case do not correspond to on-line learning episodes -- necessarily increase with development, both absolutely and as a proportion of total verb vocabulary, but the sizes of these increases do not correlate positively with children's tendency to overregularize, which is what the Vocabulary Balance hypothesis requires. It appears that something endogenous to children's grammatical systems, and not a change in either their environments or their vocabularies, causes overregularization errors to begin.

12: The Relation of Overregularization to the Development of Tense Marking of Regular and Irregular Verbs

If vocabulary development does not predict the onset of overregularization, does anything? Pinker and Prince (1988) suggested that children's lag in overregularizing might be due to the development of the regular tense marking process itself. If, consistent with traditional assumptions and contrary to the assumptions behind the Rumelhart-McClelland model, children have an ability to memorize stems and pasts independently of the ability to generate the past from the stem, and if the specifics of English tense marking take time to learn, then before they are learned correct irregular pasts could be produced. Overregularizations would be absent because the child's regular past tense marking machine is "off," not because it is "on" but starved of regular inputs.

There is nothing in this hypothesis that is incompatible with the Rumelhart-McClelland model being an accurate model of the child's past tense marking machine. One would, of course, have to give up the idea that there is no rote storage outside the past tense marking machine, and that U-shaped development can be explained entirely by the internal workings of the machine as it processes a changing input mixture. The sequence would be explained by a transition from a stage where only the rote lexicon was working to a stage at which a connectionist model began to process stems and pasts as inputs. MacWhinney and Leinbach (in press) endorse (but do not implement) such a two-module system in their connectionist model of past tense inflection. Indeed, Rumelhart and McClelland themselves briefly entertained such a possibility, and there are numerous suggestions in the connectionist literature for how to implement the equivalent of gates or on-off switches for multi-module network models (see, e.g., Jacobs, Jordan, and Barto, 1991). Thus the analyses we discuss in this chapter do not speak to how past tense marking is computed (i.e., whether it is computed by traditional symbol processing or by parallel distributed processing), as long as it can be computed independently of rote storage of lexical items.

To see if overregularization begins only when the productive regular process is first activated, we need the results of *wug* tests administered to children at various points before and after they overregularize. Obviously these data don't exist and it is not practical to hope for them. One and two year old children make poor experimental subjects in elicited language production tasks, especially with made-up words introduced on the spot. Fortunately, there are data from spontaneous speech that indirectly bear on the hypothesis.

In principle, a rule-less child with excellent rote memory could produce regular and irregular pasts with equal facility, as long as they were available in the input to be memorized. In practice, however, the young child produces correct irregular past tense forms more reliably in obligatory past tense contexts than regular past tense forms (Brown, 1973; de Villiers and de Villiers, 1973; see also Note 19). Plausibly, irregulars are easier to memorize and produce because of their higher token frequency, their phonological simplicity (all irregulars have monosyllabic roots, but many regulars do not; Pinker & Prince, 1988), their greater phonological salience (vowel changes might be more perceptible than a word-final *t* or *d*, especially as part of a consonant cluster), or some combination of these factors. However, once the regular process is acquired, the irregular advantage should be nullified: any verb, whether its past tense form is remembered or not, can be supplied with a regular past tense form at that very moment. Therefore the initiation of the regular process should be visible in spontaneous speech in a reduction or elimination of the difference between the ability to supply the past tense forms of regular and irregular verbs in obligatory past tense contexts. One can then see whether this equalization occurs near the onset of overregularization.

A related signature of the acquisition of a productive regular process can be sought in the absolute rate of marking tense on regular verbs. In English syntax all main clauses must be marked for tense. Thus we can think of syntactic processes as issuing a subroutine call to morphology, demanding a tensed form of the particular verb to be used. English, of course, executes this subroutine in different ways: the various irregular forms, and the regular process. The beauty of the regular process is that it can always keep the syntax satisfied: any verb, familiar or unfamiliar, can be suffixed to mark tense and hence can be used in a main clause (except possibly in the unusual circumstances cases discussed in Section 8). However, for a child who has not yet acquired the regular suffixation process, if the regular past form of a verb had not been previously memorized, the desire to mark tense on it would have to go unsatisfied. Mastery of the regular process, then, should be visible in attainment of high absolute rates of marking the past tense on regular verbs in obligatory past tense contexts. Again, one can see whether attainment of such levels occurs near the time of the first overregularization, and whether degree of successful regular marking correlates with overregularization across time and children. The prediction depends on the assumption that the syntactic requirement of tense marking has itself been acquired (if children do not know that tense is obligatory, they could leave regular verbs in past tense contexts unmarked even if they knew how to mark them), so if children do not reliably tense on either regular verbs or irregular verbs (which do not depend on a productive process) in obligatory contexts, this prediction cannot be tested. But if regular verbs are reliably marked, we have indirect evidence for a productive regular process; if irregular verbs are reliably marked but regular verbs are not, we have indirect evidence that the regular process has not yet developed. (The acquisition of obligatory marking of irregular verbs will be discussed in a separate section.)

Note that these predictions about levels of irregular and regular marking appear to contrast with those of the Rumelhart-McClelland model under the assumption that the model is responsible for early correct performance with irregulars, with no separate lexical storage. During the early period of learning, before the influx of regular verbs triggers overregularization, the model rapidly attained successful performance with both irregular *and* regular verbs (80-85% of correct features generated), with, if anything, a slight advantage for the regulars (see their Figure 4, reproduced here as Figure 16). Furthermore, Rumelhart and McClelland point out that their model is not just doing the equivalent of rote memorization before it overregularizes, but is showing "substantial generalization." Newly presented regular verbs are inflected with 75% accuracy (chance is 50%) on their very first exposure, based on the training during the period in which irregulars are being produced correctly.

The Development of Regular Tense Marking and Overregularization

In this section we examine whether there is an association between overregularization and the development of the productive regular affixation process. We begin with Adam, Eve, and Sarah, who show initial overregularization-free months in their longitudinal samples, allowing tests of whether the onset of overregularization coincides with development of high levels of marking regular verbs in obligatory past tense contexts (high both in absolute terms, and in terms of a rate comparable to the rate of tensing irregular verbs).

The Brown Children

We present Cazden's data on rates of supplying correct past tense forms of irregular and regular verbs in obligatory past tense contexts in Figures 33-38, and in Appendices -.[37]

Figures 33-38 here
----

[37]The data are grouped into months to be commensurable with our overregularization data. Sarah's recording sessions were half as long but twice as frequent as those of Adam and Eve; to facilitate comparisons among the three children, Cazden pooled successive pairs of Sarah's samples. Sometimes these pairs straddled a boundary between months of chronological age, the units we have been using. When this happened we divided the tokens in that pair of samples by 2, and assigned one half of them to the month preceding the boundary, the other half to the month following it.

As Brown (1973) pointed out, curves plotting use of a morpheme in obligatory contexts as a function of age are invariably noisy. In particular, early points often represent a tiny number of instances, because the grammatical structures that allows obligatory contexts to be recognized themselves develop with age, and this can result in severe sampling error (Brown, 1973). For example, the early spikes where Sarah anomalously marks irregular and regular verbs 100% of the time represent 2/2 and 3/3 tokens, respectively.

Nonetheless, several general qualitative features of the data are noticeable. First, for all three children the first month with an overregularization displayed a high level of marking of irregular past tense: 84% for Adam, 93% for Eve, and 70% for Sarah (though the neighboring months are higher for Sarah, and much lower for Eve). Second, regular marking is low before the first overregularization, but displays rapid increases to high levels, overlapping those seen for irregular verbs, shortly afterward. Adam's first overregularization occurred during a 3-month period in which regular marking increased from 0% to 100%; Eve's occurred during a 7-month period in which she went from 0% to 95%; Sarah's occurred during a 4-month interval in which she went from 0% to 78%. (Moreover for Adam and Sarah there appears to be a brief decline in regular marking coinciding with a reversion to several months of no overregularizations immediately after the first recorded one, followed by a rise to levels close to 100% around the time when overregularization resumes.)

Table 7 here

Table 7 confirms that rates of regular marking are low and smaller than those for irregular marking for the period before the first overregularization, and that the difference is narrowed or eliminated for the period beginning with the first overregularization. The increase in both kinds of marking and the relatively larger increase for regular marking can be captured in a 2 X 2 Analysis of Variance whose factors are Regular versus Irregular verbs and Before versus During overregularization, and whose dependent variable is the proportion of verb tokens marked for tense in obligatory contexts for each of the three children. Verbs were marked more reliably after overgeneralization began than before, $F(1,2) = 199.80$, $p < .005$, and, marginally, the increase was larger for regular than for irregular verbs (for the interaction between the factors of Regular/Irregular and Before/During, $F(1,2) = 8.77$, $p < .10$). The correlation between tense marking and overregularization can also be shown in continuous measures: there are positive correlations over months between the proportion of regular verbs marked in obligatory past tense contexts and the overregularization rate: $r(14) = .33$ for Adam; $r(8) = .48$ for Eve, $r(22) = .44$ for Sarah. Correlations with irregular marking rate are positive, though smaller, for all three children: .29, .42, and .21, respectively.[38]

How might one examine the second prediction of the hypothesis that overregularization is associated with the acquisition of a productive regular past tense marking process, namely, that children mark regular verbs "obligatorily" during the overregularization period? Of course literally "obligatory" means nothing less than 100% marking, which none of the children attained. Brown (1973) employed the criterion that a morpheme was being marked "obligatorily" if it was supplied in 90% of its obligatory contexts in six successive hours of speech. However, he was not interested in obligatory marking for any particular morpheme, but in the relative order of acquisition of 14 heterogeneous morphemes; the reason he adopted the 90%/6-hours criterion was because it could be used to derive an objective rank ordering of the age of acquisition of the morphemes. For our purposes, which do not involve ordering heterogeneous morphemes,

---

[38]Sarah's correlation with regular marking would be statistically significant if one were to treat months as independent sampling units; the other correlation coefficients would not be. Such significance tests, however, are meaningless because monthly transcript sets are not independent samples from a population, as noted earlier.

the criterion is in some ways too strict -- because extended use of a morpheme in (say) 75-85% of obligatory contexts clearly indicates relevant knowledge -- and in others, too lax -- because a span with a lucky streak of correct usages could be counted as mastery, only to return to lower levels because of regression to the mean. Brown notes that the developmental curves show such patterns in several instances.

One convenient benchmark is whether a child marks tense more often than not in obligatory contexts (i.e., if significantly more than 50% of verb tokens are marked). The 50% figure is, of course, partly arbitrary, but it does literally reflect a systematic preference that tense be marked. Moreover, because children initially mark most morphemes at rates far less than 50% (most notably, the regular past tense), it is unlikely to represent the base rate for the proportion of times a past tense is called for on pragmatic communicative grounds, and attainment of a level of marking greater than 50% represents movement toward the adult state where tense marking is truly obligatory.

In Table 7, marking rates that are significantly greater than .5 by a two-tailed binomial test (computed by the approximation to the normal distribution) are indicated with an asterisk. All three children, before they began to overregularize, left regular verbs unmarked more often than they marked them (Adam and Sarah marked irregular verbs more often than they left them unmarked). After overregularization had begun, all three children marked both irregular verbs and regular verbs more often than they left them unmarked. In sum, for Adam and Sarah, and to a lesser extent, Eve, there is evidence that the onset of overregularization is temporally associated with attainment of high rates of marking tense on regular verbs, comparable to rates of tensing irregular verbs.

The data from Adam, Eve, and Sarah help resolve a paradox noted long ago in the developmental literature. Ervin (1964) remarked that for some of her subjects, there were no regular past tense forms preceding the first overregularizations. Among the 9 children with extended longitudinal transcripts in our sample, we found this to be true for Naomi and

April as well: the first regular verbs marked for past tense appeared a few days *after* their first overregularization. The puzzle arises because no specific rule of morphology can be innate. Therefore the child must acquire any particular rule on the basis of individual regularly-inflected forms memorized from parental speech. All children, then, should be capable of producing at least some regular past tense forms before their first overregularization. The paradox disappears when we see that Adam, Eve, and Sarah used regular past forms at very low rates before their first overregularization, presumably because memorizing them or retrieving them from memory is difficult (though not impossible) owing to their low relatively low frequency and salience. For some children, then, the rates can be as low as zero, or at least low enough that no examples turn up in their transcripts. At some point the child extracts a suffixation process capable of generating regulars freely, apparently from a small amount of evidence if the frequency of producing regular forms reflects the number serving as the basis for the rule, and uses it simultaneously to inflect regular verbs and occasionally to overregularize irregular verbs.

The Kuczaj Children

Let us turn to Abe. As mentioned, before-and-after comparisons are not possible for him, because he overregularized in his first sample. However, Kuczaj (1977a) noted that in the period ending at 2;6 Abe's overregularization rate, and his rate of marking regular and irregular verbs in obligatory contexts, were both lower than they were in subsequent months. Figures 39 and 40, taken from the Appendices to Kuczaj (1976) (reproduced in Appendix), show the data in full.[39] High rates of regular marking (often higher than the rate of irregular marking) characterize the entire period, starting at 76%, increasing in 4 months to 98% and staying close to that level thereafter. Because Abe's data, unlike those

[39]These percentages were taken directly from Kuczaj (1976), rather than being calculated from the transcripts or raw data tables, so in some cases they are based on slightly different token counts for number of correct irregulars than the ones we have been using (i.e., the data listed in the second column of Appendix).

from Adam, Eve, and Sarah, include the later phase where overregularization begins to diminish (the right hand arm of the U), the correlation coefficient between regular marking and overregularization calculated over the entire period is low (.09). The reason is that although overregularization is lower in some of the very early months, when regular marking is also low, overregularization becomes low again in the latest months, when regular marking, of course, remains high. When the late reduction of overregularization is held constant by partialing out age, the correlation between overregularization and regular marking is far higher, .44.

Figures 39 and 40 here

As Kuczaj (1977a) points out, his cross-sectional sample shows a similar distribution to Abe's longitudinal sequence. All 14 children overregularize, and the sample includes the late span when overregularizations decrease: the three oldest children in the sample overregularize the least. All the children but one marked regular verbs in obligatory contexts at high rates (84% - 100%) except one, the youngest and lowest-MLU child, who also had the lowest overregularization rate except for the three oldest children. Statistically, the children in the cross-sectional sample behave similarly to Abe's months; though the simple correlation coefficient between overregularization and regular marking is low (-.09), the partial correlation with age held constant is high ($r = .47, p < .05$).

Summary of developmental patterns. Several general relations appear to hold consistently among the children we have examined. Before the first overregularization, regular verbs in obligatory past tense contexts are left unmarked more often than they are marked, and they are marked at lower rates than irregular verbs. In the period beginning with the month of the first overregularization, there is no consistent difference between the rates of marking regular and irregular verbs; both are marked more often than they are left unmarked, often at very high rates. During developmental spans before overregularization begins to diminish, rates of regular marking correlate positively with the rate of overregularization.

The pattern is consistent with the hypothesis that the immediate trigger for overregularization is the acquisition of the process responsible for regular tense marking, which is independent of early use of correct irregular forms. Once productive regular marking has been acquired, on any occasion that children try to mark the past tense of an irregular verb but fail to retrieve its past form, they can now inflect the stem productively. They do so for the same reason that they can now inflect regular stems at high rates, despite the initial disadvantage in memorizability that the regulars had faced.

Of course with these correlational data we do not have strong evidence that a common underlying acquisition event *causes* the correlation between overregularization and regular marking, as opposed to both phenomena simply increasing with age for independent reasons. At this point the most that can be said is that we have found consistent relations in the predicted direction between overregularization and the level of regular marking in obligatory contexts, both of them showing their major developmental increases in the same general window, before the later diminishment of overregularization decouples them. The consistent correlations in the predicted direction contrast with the results of tests of the Vocabulary Balance hypothesis in the preceding chapter, where all the predictor measures of vocabulary either remained constant or correlated in the wrong direction with overregularization.

Note, finally, that although children systematically mark tense by the time they begin to overregularize, the rates varied from child to child. It is possible that some of the individual differences among children in overregularization rates are related to these differences. Recall that Abe and Kuczaj's cross-sectional sample showed unusually high overregularization rates compared to the other children we examined. Although part of the explanation may lie in the nonrepresentative circumstances in which Abe's conversations were recorded (see Section 10), this artifact seems unlikely to account for the entire difference. Abe and Kuczaj's cross-sectional sample are also unusual in the high rate of tense marking of both irregular and regular verbs after his first samples: Abe marked 97% of his regular verbs, and the 14 children marked 93.9% of theirs, compared to Adam's 72% marking of regulars during the period studied by Cazden in which he overregularized, Eve's 66%, and Sarah's 85%.[40] For example, Abe's observed overregularization rate is 24%. Imagine that this figure comes from his failing to retrieve the past form of an irregular verb on 24.7 out of every 100 occasions in which he tries to mark it for tense, and then applying the regular suffixation process to the stem on 97% of those occasions. What would happen if Abe only suffixed verbs as successfully as Adam did? He would only have regularized the irregulars 72% of the time, resulting in 17.8 overregularizations and 6.9 stem forms. (In fact the number would be less, because Adam overregularized his non-marked irregulars in past tense contexts a smaller proportion of the time than he marked his regular verbs.) We would measure his overregularization rate as 17.8/(17.8+75.3) or 19.1%, 5 percentage points closer to the rest of the children.[41]

The Development of Irregular Tense Marking and Overregularization

---

[40]Some of the differences among the Brown children may simply reflect how old they were when Cazden stopped tallying their marking rates, and hence the proportion of early transitional samples that contributed to these means. Note as well that part of the difference between the Brown children and the Kuczaj children can be attributed to the different methods Kuczaj and Cazden used in counting correct past tense forms. Cazden apparently only counted a correct past tense form as "supplied" if it was in one of the obligatory contexts for the past tense that she used to identify no-marking errors; other correct pasts were ignored (hence the difference between the number of supplied irregular forms she tallied and the number of correct irregulars we found, as summarized in Appendices -.) Kuczaj appeared to have counted all overt past tense forms as if their very appearance was prima facie evidence that each one was obligatory in its context. This difference in method is unlikely to be responsible for much of the difference, however, because for most of Abe's months and Kuczaj's children, the estimates of tense marking were 100% or close to it; therefore, subtracting some correct past tense forms would make little or no difference.

[41]Another way of replacing Abe's rate of successful regularization with Adam's would be to use their rates of regularizing an irregular verb when the correct irregular past form was not produced (i.e., overregularizations as a proportion of overregularizations plus stem forms in past tense contexts; see Section 10), rather than their rates of suffixing regular verbs in past tense contexts. In that calculation, the discrepancy would be reduced almost entirely. The comparison would be suspect, however, because the estimate of Abe's rate of regularizing irregular tokens is not independent of his overall overregularization rate, the figure we would be adjusting.

So far, we have examined evidence on the relation between overregularization and the marking of regular verbs in obligatory contexts. The data suggest that by the time overregularization first appears, children are already marking *irregular* verbs more often than not in obligatory past tense contexts (see Table 7 and Figures 33-37 and 39; Abe's first month may be a counterexample). Presumably these early relatively high rates of irregular marking reflect partial acquisition of the syntactic requirement in English that verbs be marked for tense in main clauses, together with knowledge of the semantics of the past tense itself. Given a general motive to mark past tense on irregular verbs, the regular affixation process, once it is acquired, can serve as the means to do so when retrieval fails. Indeed, once the Brown children began to overregularize, they never did so at a rate larger than their rate of nonmarking of irregular verbs in obligatory contexts prior to overregularization. That is, when overregularizations appear they occupy some of the space in children's past tense conversations formerly occupied by bare stems. This is consistent with the hypothesis that overregularization steps in when retrieval fails; if overregularizations occurred at much higher rates than immediately preceding nonmarking rates, they would seem to represent some radical "reorganization" or global "regression," but they do not. In Kuczaj's subjects, we do not have estimates of the rate of irregular marking prior to the onset of overregularization, but even in the samples available we see that the rates of nonmarking in the youngest-age samples are far higher than overregularization rates thereafter. Indeed, acquisition of regular affixation seems to top off the reservoir of obligatory past tense contexts; as Kuczaj (1977a:593) notes, "Apparently once the child has gained stable control of the regular past tense rule, he will not allow a generic verb form to express "pastness," which eliminates errors such as *go*, *eat* and *find*, but results in errors like *goed*, *eated*, and *finded*."

The reliable marking of tense on irregular verbs, though it accompanies overregularization, does not appear to be tightly linked in time to its onset or level.[42] In individual months, children can use irregular verbs correctly in obligatory contexts more often than not without overregularizing (e.g., Adam and Sarah), and can overregularize while leaving irregulars unmarked more often than not (Abe's first month). Moreover, though the rate of irregular marking in obligatory contexts tends to correlate over months with overregularization rates in all children, the correlations were always lower than the corresponding ones between overregularization and the rate of regular marking. Nor is this entirely surprising. Although marking past tense reliably should increase the *number* of overregularization errors (and hence could correlate with the appearance of such errors for sampling reasons if the total number of irregular past forms were low, which it is not for Adam and Sarah), it does not necessarily increase the *rate* of overregularization calculated over all irregular verbs marked for past tense. That is because if a child neglected to mark irregular verbs for tense, we would see both few overregularizations and few correct irregular past tense forms (with unmarked stems taking their place), all things being equal.

Nonetheless we can speculate on two conceivable reasons why past tense overregularization does tend to occur in the presence of high rates of irregular past tense marking. One is that the acquisition of the obligatory tense requirement in English might be an impetus for the acquisition of the regular suffixation process itself, which is a means for the tense requirement to be satisfied for all verbs, memorized or not. (We return to this possibility in Section 14, where we consider how a regular rule might be learned.) Second, because tense is stored as part of the lexical entry of irregular verbs, they might tend to be activated by normal lexical retrieval processes driven by the intent to communicate certain kinds of semantic content. In comparison, activation of the regular operation, a grammatical

---

[42]Note that this is a modification of a conclusion stated in an earlier draft of this monograph circulated as a technical report, which was based on the ages at which Brown's 90%/6-hour requirement had been met, before Cazden's full data set was made available to us.

process, might be invoked more often for grammatical reasons (to satisfy the tense requirement). That is, *broke* automatically comes to mind when one is thinking simultaneously of breaking and of pastness, whereas thinking of fixing and pastness summons only the word *fix*, plus a call to the regular tense-marking operation. If lexical retrieval is more automatic or reliable than on-line application of regular marking, then once a child's language system is driven by the grammatical requirement to mark tense, there is a motive to apply regular marking when retrieval of an irregular fails, and overregularizations will result. In contrast, without the requirement, tense might only be marked when the irregular form was lexically activated for semantic reasons. Whether these speculations can be supported by independent data, or whether they are necessary at all, we leave as a question for future research.

In sum, it seems that acquisition of the regular suffixation process of English is the proximal trigger for overregularization errors: overregularizations first appear during a window in which the child goes from leaving most regular verbs unmarked in past tense contexts to marking them more often than not, and at rates comparable to those for irregular verbs. An overall tendency to mark past tense on irregular verbs, combined with imperfect retrieval of irregular forms from memory, and a regular process that is capable of applying to any stem, even if closely tied to an irregular past, sets the stage for overregularization errors.

13: Factors Causing Differences in Overregularization Rates Among Verbs

Although overregularization rates are low in general, not all verbs are overregularized at the same rate, and some verbs are overregularized by some children at some stages more often than they are produced correctly. By examining what it is about a verb that makes it more or less likely to be overregularized, one can test many hypotheses about the psychology of overregularization.

In this section, we correlate a variety of lexical factors with overregularization rates across a large set of irregular verbs. To minimize averaging artifacts, we calculated these correlations for each of the 19 ChiLDES children with individual transcripts that overregularized at least once. When we report averages for these correlations, and tests of the average against a null hypothesis of zero, we first transformed each child's correlation coefficient (r) to Fisher's z; the reported mean correlations were obtained by averaging the z-scores and transforming the mean z back to a correlation coefficient.

In addition, when suitable we computed a single measure that aggregates overregularization tendencies for verbs across all 25 children. Pooling tokens across children is inappropriate, as the children with the largest samples would dominate the means. However, averaging together each child's overregularization rate for a given verb is also not appropriate, because such means are in danger of being artifactually influenced by individual children with extreme values and idiosyncratic subsets of verbs. For example, if a child is fond of using the verb *hear* also happened to be a high overregularizer across the board, then the mean overregularization rate of *hear* across children would be artifactually high. This problem is accentuated for small samples of a given verb for a given child, which can contribute extreme high or low overregularization rates to the mean across children. To minimize these artifacts, for each child we only considered verbs that he or she used at least

10 times in the past tense. Furthermore we standardized each child's set of overregularization rates for different verbs (i.e., for each child we converted the overregularization rate for each verb to a *z*-score, using the mean and standard deviation of the overregularization rates for different verbs used by that child). The *z*-scores for a given verb were then averaged across children. These scores, which representing the relative overregularization rate for each verb in the aggregate sample, are summarized in Appendix . For the convenience of the reader we have also linearly rescaled these mean *z*-scores back to proportions that can be more intuitively interpreted as "average overregularization rates." For each verb we multiplied its *z*-score by the mean across children of the standard deviations that had been used to calculate the *z*-scores (i.e., the mean of the 25 standard deviations, one for each child, of the overregularization rates for that child's different verbs), added it to the mean overregularization rate for that verb (including 0's from the children who never overregularized), and added an additional constant of .02 to raise the least-overregularized verb from a negative value to 0. Note that these figures are primarily for ease of comparing relative overregularization rates among verbs, and do not literally correspond to a mean across children.

Frequency

As we discussed in Chapter 8, if overregularization results from a failure to retrieve a listed past tense form, forms with greater memory strength should be more resistant to overregularization (Slobin, 1971; MacWhinney, 1978; Pinker, 1984). The more often a parent uses a past tense form, the stronger the memory trace for that form should be, and the stronger the association between it and the corresponding stem form. Thus the adult frequency of an irregular verb in its past tense form should be negatively correlated with its overregularization rate for children. Bybee and Slobin (1982a) found a significant negative rank order correlation over verbs between their preschool children's overregularization rates and the frequencies of the verbs in the speech of the preschool children's caretakers.

They found a similar effect for some subclasses of verbs for the past tense forms experimentally elicited from their third grade subjects. MacWhinney (1978) documents similar findings in children acquiring other languages. We sought to replicate this effect with our larger sample of children, and to examine it within individual children, to ensure that it is not an averaging artifact.

Three measures of adult frequency were used: the adults talking to a particular child in the transcripts, an aggregate parental frequency measure computed by averaging the frequencies of the past tense forms of the verbs in the speech of the 19 sets of adults, and the past tense counts from Francis and Kucera's (1982) corpus of a million words of written text. (We include correlations with the Francis and Kucera database because it is the most commonly used source of frequency information in psycholinguistics; it should predict children's behavior less well, of course, because it is from written English addressed to adults). Frequencies and mean frequencies were converted to logs because the frequencies could range over several orders of magnitude (especially for Francis-Kucera figures) and we expected that a frequency difference of 1 versus 10 would have a greater effect than a frequency difference oi 1001 versus 1010.

Frequency had a very clear effect. The aggregate overregularization rates across 19 children significantly correlate with the aggregate parental frequency counts, $r(37) = -.37$, $p < .05$, and nonsignificantly correlate with Francis & Kucera frequencies, $r(37) = -.14$, $p > .10$. Of the 19 children, all 19 had overregularization rates across verbs that correlated negatively with the log aggregate parental measure (range -.12 to -.61); 18 correlated negatively with the log frequency of their own parents; and 16 had negative correlations with log Francis & Kucera frequency. These sets of individual children's correlation coefficients have means that are significantly less than zero: mean r -.33, $t(18) = 8.89$, $p < .001$ for the aggregate parental frequency; mean $r = -.34$, $t(18) = 6.80$, $p < .001$ for each child's own parents; mean r -.16, $t(18) = 4.45$, $p < .001$ for Francis-Kucera frequency.

Phonological Similarity Between Stem and Past

An important fact about irregular verbs is that their past tense forms, though unpredictable in other regards, generally preserve most of the phonological composition of their stems. *Go-went* and *be-was* are exceptions; for the other irregulars such as *come-came*, *feel-felt*, and *bring-brought*, the past and stem overlap to an extent that would be uncanny if the pair consisted of two arbitrary words linked only as memorized paired associates. Pinker and Prince (1988) point out that theories of irregular morphology should explain this fact, but that the Rumelhart-McClelland model failed to do so.

Number of Changes from the Stem Form to the Past Form

In some theories of generative grammar (e.g., Chomsky & Halle, 1968; Halle and Mohanan, 1985), irregular pasts are generated by applying to the stem one or more rules that replace a circumscribed substring of phonological segments. It is in the very nature of rules that any segment not changed by the rule is left untouched and hence will automatically appear in the past tense form. This is how the similarity between the members of an irregular pair is explained.

MacKay (1976) suggested that these rules are applied by speakers on-line when they produce irregular past forms, and that each application consumes a determinate portion of processing resources. When comparing the response times for adult subjects to produce past tense forms when given their stems, he found that verbs with "simple" vowel changes were produced most quickly, followed by regular verbs, followed by verbs with "complex" vowel changes, followed by verbs with both a vowel change and the *t* suffix. If MacKay's hypothesis that the psychological complexity of irregulars is predicted by the number of rule applications in the grammatical derivation, as proposed by irregular-rule theories, such effects may also affect children's overregularizations. Irregulars with more changes require more rule applications and hence may be harder to produce; when the derivation breaks down, the regular rule steps in.

This hypothesis can be tested by correlating the number of phoneme changes that must be executed to derive the past from the stem with the overregularization rate for that stem. What counts as a "change" will of course depend on one's theory of possible phonological operations, but a reasonable if crude first approximation would be to count each single vowel substitution, consonant substitution, consonant addition, or consonant deletion as one change. For example, in *see-saw*, one phoneme, the vowel, must be replaced; in *sweep-swept*, a vowel must be replaced and a consonant suffixed, for a total of two. Hence on average a verb like *sweep* should be overregularized at a higher rate than a verb like *see*. In these calculations, we treated each diphthong as a single phoneme.

We did not find a consistent positive correlation between Number of Phonemes Changed and Overregularization Rate: for the aggregate rates across the 19 children, $r(34) = -.10$.[43] For 14 of the 19 children, correlations were also negative, as was the mean of the correlations (-.08).

Degree of Phonological Overlap between Stem and Past Forms

It is possible that there really are no irregular rules, and that the commonalities between stem and past (a heterogeneous set of patterns which range from simple vowel changes to such severe distortions as *bring-brought*) are to be accounted for by other means. For example, these pairs may have been generated in earlier stages of the language by genuine rules, now defunct; the pairs that were produced by these rules, because they shared phonological material, were easy for learners to memorize, and have preferentially survived over the centuries in Darwinian fashion (see Bybee and Slobin, 1982a, b; Lieber, 1980; Pinker and Prince, 1988).

One extreme mechanism by which stem-past similarity might operate is by affecting

[43]For Adam, Eve, and Sarah, no-change verbs were included; *have*, *be*, and *do* were excluded because it is not clear how irregular rule theories would treat them (all three have irregular present tense forms, and *be* has two past forms *was* and *were*, neither related to the infinitive.)

the likelihood with which children recognize that a stem form and its past tense counterpart are alternative versions of the same verb, as opposed to two independent verbs. In fact, it has been proposed that failure to unite two inflected forms within a single inflectional paradigm for a given verb is an important general cause of overregularization, accounting for why it takes place despite the Blocking Principle (Kuczaj, 1981; Bybee and Slobin, 1982a; Pinker, 1984; Pinker and Prince, 1988; Clark, 1987, 1990; see Note 5). Two phenomena suggested this hypothesis. First, Kuczaj (1981) and Pinker & Prince (1988) noted that children productively inflect irregular stems: *ate*, *ated* and *ating* coexist with *eat*, *eated* and *eating*, as if they were two verbs. Second, Bybee and Slobin (1982a) noted that irregular verbs that end in a vowel that changes from stem to past (*fly-flew, see-saw, know-knew, blow-blew*, etc.) are particularly prone to being overregularized. If phonological overlap between stem and past is a critical cue for two forms to be lumped together as versions of the same verb, the meager common portion among verbs in this class (e.g., initial *s* for *see* and *saw*) would make it harder for the child to recognize that they were forms of a single verb.

The hypothesis that a failure to unify two stems within a common verb paradigm is the major cause of overregularization appears to be too strong. Empirically, errors like *ating* and *thoughting* are uncommon, and tend to appear later in development than simple overregularizations (Kuczaj, 1977a, 1981). Theoretically, the cues that would tell the child that the two forms were versions of the same verb are present throughout development, leaving it a puzzle that the child takes so long to notice them. Moreover the linguistically valid cues for common membership militate strongly against the child's ever considering the past stem to be an independent verb. Kuczaj (1981) notes that the child correctly treats the past stem as indicating pastness: *was wenting* occurs; *is wenting* does not. But according to crosslinguistic research in lexical semantics, tense is an extremely unnatural, perhaps nonexistent, inherent semantic component of verbs (Bybee, 1985; Talmy, 1985; Pinker, 1989): Languages do not like independent verbs that mean "do X" and "do X in the past." If children's hypotheses mesh with what is linguistically possible, they should not posit such verbs.

Weaker versions of the hypothesis are plausible, however. Perhaps phonological dissimilarity between stem and past does not have so large an effect that it prevents children from mentally linking them over the long term, accounting for the very existence of overregularization, but has a smaller effect over the short term, causing slight delays in mentally unifying past and stem for verbs with dissimilar pasts compared to verbs with similar pasts. Moreover, if links between stem and past are not immediately acquired at full strength, stem-past similarity could have an influence on children's ability to retrieve the past form after some kind of link had been formed. This could come about if on some occasions the child retrieves the past only after first mentally activating the stem. Common phonological material between stem and past might improve the speed or probability with which a past form is activated, if, for example, both forms are linked via pointers to a common representation of their phonological constituents, or if they actually overlap in their underlying phonological representations at some level. That is, it might be easier to retrieve *strung* given *string* than *brought* given *bring*.

If phonological similarity has an effect on overregularization, it would not be the degree of phonological change, but the degree of phonological overlap, that should be related to overregularization-proneness. The degree of phonological overlap for irregular pairs was quantified as follows. We began by counting the number of phonemes preserved from stem to past. For example, *forget-forgot* has 5 phonemes preserved, whereas *catch-caught* has only 1 preserved. Counting shared phonemes alone would be highly misleading, however, because speakers certainly represent words in a format that is more structured than a simple list of segments. In particular, the consonant-vowel (CV) skeleton underlying a form is thought to be a distinct level of representation in phonology (see Kaye, 1990), and this would saliently capture the relatedness of *see* and *saw* or *throw* and *threw* even if the sheer number of shared segments was slight -- the *lack* of a consonant at the end of a verb is *itself* a feature of similarity. Thus we also counted the segments that changed content but preserved their positions in the verb's CV skeleton (with half the weight of phonemes that preserved both their content and their skeleton position), so that *see* and *saw* were counted as more similar than *say* and *said.* Similarly, consonants that preserved their content but changed position in the CV skeleton were given a quarter as much weight as consonants that preserved both content and position, so that *l* was given less weight in *feel-felt* than in *steal-stole*). (As it turns out, nothing hinges on this particular way of measuring stem-past similarity; similar results occur if all phonemes are weighted equally, or if shared phonological features, rather than shared phonemes, are counted.)

The prediction is that the amount of preserved phonological structure should correlate negatively with overregularization rates; that is, the more similar the stem and past, the less likely the child is to overregularize the verb. At first it might appear that the prediction has already been tested in the preceding section: every phoneme that is changed is a phoneme that is not preserved, so the number of phonemes changed (previously tested) should correlate negatively with the amount of overlapping phonological structure (to be tested here). However, the correlation is only -.55, leaving enough independent variance to test for a negative correlation with overregularization rates even though we just saw that the phonemes-changed measure had no effect.

In fact, the correlation is positive for the aggregate overregularization rate (.10), for the mean correlation for these children (.07), and for 14 of the 19 children individually.

One might note that the irregular with the least phonological material preserved, *go-went*, is high in frequency. This suggests that any difficulty in learning more dissimilar past pairs might be compensated for by the higher frequency of such items, masking an effect of dissimilarity per se. Indeed, the correlation between aggregate parental frequency and phonological overlap is - 31 ($p < .005$). An effect of phonological overlap could be unmasked in a multiple regression analysis using Frequency and Phonological Overlap as predictors. Such an analysis fails to reveal an effect of phonological overlap: the partial correlation, holding constant the child's own parental frequency, is in the nonpredicted direction for the aggregate measure over children and for 10 of the 19 individual children. (Similar non-effects are found when frequency is partialed out of the correlation between overregularization and the number of phonemes changed, which is not surprising, because the correlation between the two predictors is only -.05). In sum, we have failed to find clear evidence that the degree of stem-past similarity or dissimilarity affects the overregularization rate of irregular verbs.

The failure of any measure of relative stem-past similarity to explain differing overregularization rates is unexpected given the pervasiveness of such similarity in the English irregular verb system. Perhaps a modicum of phonological similarity, together with the obvious semantic similarity between stem and past, may be sufficient for children to recognize the relatedness of most verbs and their irregular pasts. Thereafter, irregular pasts are retrieved directly in response to a call to the entire lexical paradigm and the "past" feature, bypassing the stem form itself (as was suggested in Sections 10 and 10).

Protection by Families of Similar Irregular Pairs

Not only are irregular verb stems similar to their past tense .orms; they tend to be similar to other irregular verb stems that have comparable past tense forms. Irregular verbs fall into clusters such as *sting-stung*, *swing-swung*, *string-strung*, and so on. The minor rules posited by some theories of generative morphology (e.g., a rule changing *i* to ^; see Halle & Mohanan, 1985) are meant to explain this second kind of similarity as well. The form of irregulars is by definition unpredictable on phonological grounds, so the rules must be tagged as applicable only to a fixed list of words, but if the number of rules is smaller than the number of words, the existence of similarity clusters is explained.

However, Pinker & Prince (1988) pointed out problems for such theories. The irregular clusters are held together by far more common features than just the segment changed by the putative rule: *string*, *sting*, and *swing* share not only an *i*, but a velar nasal as their final consonant and an *s* as part of an initial consonant cluster. Trying to capture these *hypersimilarities* by adding them to the rule as context terms (e.g., "change *i* to ^ in the context *C_ng*") fails in both directions. It falsely includes many forms like *bring-brought* and *sing-sang*, and fails to include verbs that are clearly related to the cluster by family resemblance, such as *stick* (final consonant velar but non-nasal) and *spin* (final consonant nasal but non-velar); see also Bybee and Slobin (1982a).

In some ways the Rumelhart-McClelland model handles these imperfect partial similarities well: after being trained on 82 irregulars, some of the model's outputs for new irregulars it had not previously encountered were correct, such as *wept*, *clung*, and *bid*, despite the complex and highly probabilistic nature of the patterns that such generalizations represent. Furthermore the model proved to be highly sensitive to the subregularity that no-change verbs all end in a *t* or *d*, overgeneralizing it to regulars and to other irregulars that end in *t* or *d*. This can be attributed to the fact that the model records the relative frequencies of many different mappings between substrings of stems and substrings of pasts and it superimposes them across the different verbs that exemplified them.

Pinker and Prince (1988), while disagreeing with Rumelhart and McClelland's suggestion that both regular and irregular forms are generated in a single associative network, noted that their model might offer insights as to how irregular verbs are stored. If the traditional notion of rote memory for irregular storage is thought of not as an unstructured list of slots, but as also involving some kind of associative network in which recurring similarities are recorded and superimposed, the hypersimilar family resemblance classes can be explained because they contain sets of verbs that are easier to memorize than unrelated singletons, and are prone to occasional generalizations (e.g., *brung, bote*) by analogy. In this interpretation Rumelhart and McClelland would be providing a better model of the irregular rote component of the inflectional system. Related suggestions had been made prior to Rumelhart and McClelland (1986) by MacWhinney (1978) and Bybee and Slobin (1982a). MacWhinney suggests that some children utter some productive forms by a mechanism distinct both from rote lexical storage and from rule application, namely rhyme-driven analogies. Bybee and Slobin suggested that speakers form *schemas* for recognizing typical phonological patterns of irregular past tense forms. Children learn to associate past forms with their stems more easily if they conform to a past tense schema, and they are more likely to select stored forms that conform to a schema when producing past tenses.

If belonging to a family of similar irregulars undergoing similar changes strengthens the memory trace of a given irregular form, it should be more resistant to overregularization than more isolated irregulars, holding frequency constant. The prediction that partial regularity blocks overregularization was first suggested by Slobin (1971), and has been further tested by Kuczaj (1977a, 1978) and Bybee and Slobin (1982a). The most robust effect is that verbs that end in *t* or *d* are less likely to have *-ed* added, and are more likely to be uttered in no-change form, than verbs without those endings. This is true both for no-change irregulars, leading to improved performance, and for other kinds of irregulars, leading to no-change errors; both kinds are protected from overregularization (see Rumelhart & McClelland, 1986, and Pinker & Prince, 1988, for reviews.) (We see the effect in our data as well: Of the 11 verbs ending in *t* or *d* listed in Appendix , 10 of them were overregularized at aggregate rates lower than the mean across verbs.) As mentioned, the Rumelhart-McClelland model duplicated this phenomenon. However, Pinker and Prince (1988) point out that the effect is potentially so overdetermined that identifying the psychologically active cause or causes is nearly impossible. The no-change class is large (the largest among the irregular verbs), shows an exceptionless hypersimilarity (all its verbs end in *t* or *d*), involves a single kind of change (none), shares its verb-final consonant with the regulars, and when regularized results in a phonological pattern (adjacent identical stop consonants) that the phonology of English tries to avoid. Thus the existence of an effect of family strength should be confirmed with other materials.

Bybee and Slobin also showed that children overregularized subclasses with different kinds of vowel changes at different rates. They attributed the differences to different degrees of stem-past similarity, as discussed in the preceding section. But as we have seen, the effects of stem-past similarity are difficult to demonstrate if they exist at all. Furthermore, Pinker and Prince (1988) showed that stem-stem similarity may be the more relevant factor: the overregularization rates for the different vowel change classes correlate well with the number of English irregular verbs sharing the vowel changes that the class members undergo. They suggested that this explains why the Rumelhart-McClelland model mimicked the ranking of overregularization rates for these subclasses, at least in one stage. However, even here the existence of a stem family effect was not perfectly clear: Bybee and Slobin's subclasses, as interpreted by Rumelhart and McClelland, were heterogeneous and contained many possible contaminants, such as the inclusion of unusual *go-went* in the *blow/grow/know* subclass (see also Egedi and Sproat, 1991, for discussion).

A better test of the family strength effect would eschew the necessarily imperfectly constructed subclasses in favor of a direct measure of the strength of the family members for each irregular verb. Because there are many different possible models of associative memory (differing in how widely their associations spread in phonological space), none of them simulated perfectly by any given family strength measure, we computed three different measures, of differing degrees of inclusiveness, all of them based on the principal dimensions of similarity within families of irregulars (Pinker & Prince, 1988). The first measure was based on rhymes: for each verb we summed the frequency (not the log frequency) of the past tense forms of each of the other irregular verbs whose stems and past tense forms rhyme with those of the verb in question. For example, for *sting-stung* we would aud the frequencies of *clung*, *flung*, *swung*, and so on. The verb's own frequency was not included; although it surely affects the strength of the family it belongs to, we wanted to see if we could find independent support for a family strength effect, unconfounded by the frequency effect already documented. The second measure was based on the final consonant cluster: we summed the frequencies of the past tense forms of all the irregular verbs that shared the final consonant cluster with the verb stem in question and that underwent the same change from stem to past (vowel change, consonant change, consonant addition, and so on; see the subclasses in the Appendix to Pinker & Prince, 1988). For example, for *stick-stuck* we would add the frequencies of *struck* and *snuck*, even though their stems, *strike* and *sneak*, do not rhyme with *stick*. The third and most inclusive measure added the frequencies of the irregular verbs that shared a final consonant with the verb stem in question, and that underwent the same change from stem to past. For example, for *stick-stuck* we would add the frequencies not only of *struck* and *snuck*, but also of *stunk* and *slunk*. Verbs ending in a vowel were treated as if they shared a final consonant.

Because many irregular neighbors like *slunk* would be far-fetched candidates for children's lexicons, we actually selected members of irregular word families, and took their frequencies, from the adults' speech in the transcripts of the child in question, using the 19 children with individual transcripts that overregularized at least once. The independent variables took on very different values for each child, so aggregate measures are not appropriate.

The prediction we are testing is as follows: the higher the frequencies of an irregular verbs' family members, the less likely the verb is to be overregularized. The prediction appears to be borne out. For the family of rhymes, the correlation coefficient between family strength and overregularization rate was negative for 17 of the 19 children. The mean of the correlations, -.07, is significantly different from zero, $t(18) = 2.23$, $p < .05$. For the family of verbs sharing a final consonant cluster and a past tense change, the correlation was negative for all 19 children, with a mean of -.11, $t(18) = 9.99$, $p < .001$). And for the most inclusive family, sharing a final consonant and a past tense change, the correlations were also negative for all 19 children (mean -.11, $t(18) = 10.03$, $p < .001$.

To ensure that this effect cannot be attributed to a confound with the frequency of each verb, we held the log parental frequency constant in a partial correlation analysis. For all three family sizes, 16 of the 19 children had negative partial correlation coefficients, and the mean partial correlation coefficient remained negative and significantly less than zero (for rhyme, mean $r = -.08$, $t(18) = 2.22$, $p < .05$; for final consonant cluster, mean $r = -.08$, $t(18) = 4.005$, $p = .001$; for final consonant, mean $r = -.08$, $t(18) = 4.01$, $p = .001$.)

We conclude that there is a small though reliable effect whereby verbs are protected from overregularization to the extent that they are phonologically similar to other verbs (weighted by their frequencies) displaying the same irregular pattern.

Attraction to Families of Similar Regular Verbs

The preceding analysis confirms the hypothesis of Slobin, Bybee, and Kuczaj that partial regularity blocks overregularization. It is consistent both with the Rumelhart-McClelland model and with Pinker and Prince's augmentation of the traditional rote-rule model in which the rote component has some associative-memory-like properties. A test that might distinguish the latter two models is whether families of similar *regular* verbs pull an irregular *toward* overregularization, in the same way that families of similar irregulars pull it away. Since in the traditional rote & rule model regular past tense forms need not be stored because they can be generated by a rule, under the simplest hypothesis regular past tense forms should not attract irregulars. Of course, storage of regulars is possible within such theories -- what could prevent it, given that it allows the learning of individual irregular items? -- but such storage is not necessary except under certain circumstances in which the existence or form of the regular is not predictable, such as in the child before the regular rule is learned, and doublets such as *dived* and *dove* where both members must be stored. Thus in general, we should not find strong effects of storage of regulars in a rule theory, and in no case should the ability to generalize to new forms depend on the previous storage of similar old ones. That this property does in fact distinguish rule-based theories from the Rumelhart-McClelland model was shown by the behavior of the model on newly presented regular verbs. At asymptote the model erred on 33% of the regular verbs it was tested on, producing no output at all for six that were dissimilar from those in its training set, such as *jump* and *pump* (Pinker and Prince, 1988; Prasada and Pinker, 1991). In contrast, adults easily regularize highly unusual sounding novel forms such as *ploamph* or *keelth* (Prasada and Pinker, 1991), and the 2-5 year old Abe left virtually no regular verb in past tense contexts unmarked, including his own unusual inventions *eat lunched*, *bonked*, *borned*, *axed*, *fisted*, and *poonked* (Kuczaj, 1977a).

To test whether families of regular verbs pull similar irregulars toward overregularization, we first extracted the 1826 regular verbs rhyming with the irregulars that were listed in an on-line version of Webster's 7th Collegiate Dictionary, which contains phonological representations of 8217 verbs. Just as for the irregular families, three different size nets were cast. In the first, the sum of the frequencies of the regulars in a parent's speech that rhymed with each of the child's irregulars was computed (for example, *winked* would contribute to the regular attractor strength of *stink*). In the second, verbs with the same final consonant cluster were grouped, so that *yanked* and *honked* as well as *winked* join the family of *stink*. In the third, verbs with the same final consonant were grouped, so that *hiked* and *harked* would also belong to *stink*'s attracting family. Verbs that ended in a vowel were treated as if they had a common final consonant.

If families of regular verbs pull similar irregulars toward overregularization, correlations between regular family strength and overregularization rate should be positive. In none of our tests do we find statistically significant correlations in that direction. For families of rhymes, the correlation was positive for 7 of the 19 children, with a mean of -.01. For families sharing a final consonant cluster, correlation coefficients were positive for 11 of the 19 children, with a mean of .08, not significantly different from 0, $t(18) = 1.14$, $p > .10$. For families sharing a final consonant, the correlation was positive for 11 of the 19 children, with a mean (.06) not significantly different from 0, $t < 1$. When the verb's own frequency is partialed out, the correlations with the smallest family strength measure (rhymes) are positive for only 8 of the 19 children and the mean correlation across children was negative. For the middle-sized family (shared final consonant cluster), the correlations were positive for 12 of the children; the mean correlation was .09, not significantly different from zero, $t(18) = 1.45$, $p > .10$. For the largest (final consonant) family, the correlations with family strength were positive for 12 of the 19 children, with a mean of .06, not significantly different from zero ($t(18) = 1.11$, $p > .10$). Furthermore, it is not the case that the negative correlations come from children with smaller (hence noisier) samples; of the 7 children who went in the nonpredicted direction, 6 were among the 14 children whose samples contained more than 100 irregular past tense tokens, and one of them was Abe, who had the richest database of overregularizations.

Each of these measures were computed a second time, excluding polysyllabic verbs. Pinker and Prince (1988) point out that all irregulars are monosyllabic except for forms like *forget*, *understand*, and *overthrow* that contain a prefixed monosyllabic irregular. Therefore if a verb is nonprefixed and polysyllabic, it is guaranteed to be regular. If a child became sensitive to this contingency, he or she could sequester such verbs from the mechanism giving rise to family strength effects, and our including them in our estimates could dilute our estimates of such effects. But the results were virtually identical in all cases.

In sum, correlations between children's overregularization rates and various properties of different irregular verbs yield the following conclusions. As expected, irregular verbs with lower frequency past tense forms are more likely to be overregularized, underscoring the important role of the memory strength of the irregular past tense form in the overregularization process. The degree of phonological similarity or dissimilarity between a stem and its past tense form appears to have no influence on overregularization, suggesting that the errors are not primarily caused by difficulties in executing phonological changes in generating irregular forms, nor in uniting stems and pasts as part of a single verb paradigm during learning. We found evidence that irregular verbs are protected from overregularization by families of similar irregulars, though we failed to find evidence that they are drawn toward overregularization by families of similar regulars. This supports Rumelhart and McClelland's assumption that irregular patterns are stored in an associative memory, but fails to support their assumption that the regular pattern is stored in the same system.

14: Summary and Conclusions

The facts of overregularization can be summed up simply. After a period in which all of the child's past tense forms of irregular verbs are correct, the child begins to overregularize. Overregularization then occurs at a low rate throughout the preschool and early school years, affecting all irregular verbs, to an extent that depends on the verb's rarity in parental speech. Its overall rate appears to be independent of changes in the mixture of regular and irregular verbs in the child's speech, the child's parents' speech, or the child's vocabulary. Instead it seems to depend on the acquisition of the tense marking system as a whole: development of the ability to mark regular verbs reliably for tense appears to be the immediate harbinger of overregularization, and reliable marking of irregular verbs for tense accompanies it. Aside from frequency, verbs' proneness to overregularization depends to a small extent on the strength of the verb's phonological neighborhood: clusters of similar irregular verbs protect one another from overregularization. In contrast, clusters of similar regular verbs do not appear to pull an irregular toward overregularization.

These facts can be accounted for by a simple theory. The child stores irregular past tense forms in a rote memory system, in which the strength of a memory trace is monotonically related to the frequency with which it is encountered. In addition, this memory system has some of the properties of an associative network: stems-past pairs displaying similar relations reinforce each other. (This same property occasionally leads to irregular generalizations such as *brang* and *wope*.) Regular past tense forms, in contrast, are generated by a mental concatenation operation that affixes a suffix to a stem. Because this rule can always be applied on-line, regularly inflected forms need not, in general, be stored (though they can be under certain circumstances, such as in children before they have learned the rule). And because it simply adds an affix to the end of a stem with unspecified properties, the similarity of a given stem to previously encountered ones plays no role.[44] The two systems interact in a simple way: the retrieval of a stored irregular entry blocks the application of the regular rule.

The fact that overregularizations are a small minority of irregular past tense utterances at all stages shows that the Blocking process is active in the child as soon as there is evidence for two modes of inflection at all. When overregularizations do occur, they are straightforwardly explained as a failure to retrieve the irregular past form (or, for past+*ed* errors, its "past" feature) in real time. This tendency is related to the frequency of the form in an obvious way that is an immediate consequence of the logic of irregularity and the fallibility of human memory. In the extreme case, an irregular form that has been attended to with zero frequency (e.g., *shend* for adults, and many irregular verbs for young children) will have no memory trace and hence will be retrieved from memory with zero probability, and will always be overregularized if the form is to be tense-marked at all. An irregular form that has been heard once has a weak memory trace and hence a probability of being retrieved that is greater than zero but less than one. Irregulars that have been heard more times have correspondingly stronger memory traces and lesser overregularization probabilities; irregulars that have been heard thousands of times will be successfully retrieved virtually always. The learnability problem of recovering from errors is solved by a Blocking principle that operates throughout development, fed by irregular forms whose potency increases with increasing exposures during development.

Of course there are aspects of overregularization that remain to be explained. There is considerable unexplained variance in exactly which verbs are overregularized at which rates and ages, and among different children, and the temporal relation between obligatory

[44]The variation in the phonetic form of the regular affix, whereby *d* surfaces as *t* following unvoiced consonants and *id* following coronal stops, are attributable to general phonological processes operating throughout English, not to the regular process itself; see Pinker and Prince, 1988).

tense marking and overregularization is as yet unclear. At present we believe we have run up against the limitations of the available data. Just as prior characterizations of overregularization were unclear or misleading owing to the limitations of paper-based diaries and transcript samples, current computer databases of ChiLDES size, though they have enabled tremendous progress, are still not up to the enormous task facing us in trying to understand details of language development. An hour of speech a week is still a paltry sampling of the richness of the language acquisition process, especially in the critical third year of life. We suspect that child language databases with sampling frequencies an order of magnitude greater than current ones would resolve many of uncertainties in the current conclusions.

Observations on the Nature and Learnability of the Tense Marking System

Our suggested explanation of the overregularization process has three parts: decision to mark tense, imperfect memory retrieval, and possession of a regular process with universal applicability. Though the second postulate reflects a possibly uninteresting memory limitation of the child, the other two reflect quite remarkable linguistic accomplishments.

Observations on Tense Marking

As Slobin (1982) and Pinker (1982, 1984) point out, obligatory grammatical constraints pose difficult learning problems for the child. The fact that an inflection is obligatory means that there are no pragmatic cues to the semantic features that the inflection is encoding; parents must express the pastness of a past event, regardless of how relevant it is in the conversational context. Moreover once the child somehow has figured out that past tense inflection encodes past tense, if he or she mistakenly assumes that it is optional, no parental input short of negative evidence can contradict the assumption. Slobin and Pinker thus suggest that the child is innately prepared to consider obligatory inflectional tense-marking as a possible constraint in the language to be acquired; such a hypothesis is easily disconfirmed in systems where inflection for tense is in fact optional or not available. Our seemingly homely explanation for overregularization in terms of retrieval failure depends on the child having solved these daunting learnability problems (see Pinker, 1984, for an explicit hypothesis as to how the child solves them, and Stromswold, 1990, for evidence that children solve them quickly; the problems are of course finessed in network simulations that are fed correct stem-past pairs in isolation). If the 2- to 3-year-old child did not consider tense marking to be obligatory, the failure to retrieve an irregular past in a past tense context would not automatically lead to overregularization; the child could simply leave the verb unmarked.

Observations on Regular Rules

A second noteworthy linguistic achievement of the late 2-year-old is possession of a process that is capable of yielding an inflected output form for any verb, no matter how strongly linked with an idiosyncratic irregular form, and regardless of whether a family of similar regular forms is available to serve as an analogy-supporting model. As Pinker and Prince (1988) point out, a rule that simply concatenates an affix with a stem, characterized in terms of a variable standing for any stem, rather than particular patterns of the typical phonological contents of stems, easily provides this capability. A set of associations between stem phonology and past tense phonology is tied to varying degrees to the patterns it has been trained on. While such models could, in principle, approximate the unlimited applicability of a rule by training it on a set of regular stems that span enough of the phonological space of English to cover all cases, in practice this ability is compromised by the necessity of curbing links to the regular ending in order to avoid application to the irregulars. Presumably it is for this reason that the Rumelhart-McClelland model failed to display the appropriate generalization abilities for novel regulars (Prasada and Pinker, 1991). In managing to come out with a past tense form close to 100% of the time when called for, despite less than perfect memory retrieval, the children we have studied (most notably Abe) clearly have mastered an inflectional process of very wide applicability. If this regular process was not capable of applying to arbitrary irregular stems (and in fact 90% of Abe's irregulars were regularized at least once, together with many unusual creative forms like *poonked*), or if it depended on the existence of similar high frequency regulars (which it does not seem to), then when faced with irregular retrieval failure the child would be left with no choice but to utter the stem, even if his language system had called for a past tense form.

We envision regular rules as mental symbol concatenation operations, similar in operation to the core rules of syntax (see Pinker & Prince, 1988, 1991; Prince and Pinker, 1989; Pinker, 1991). This hypothesis is about a kind of mental machinery available to all language learners but put to use in different ways in different language systems; in English it is used in a condition-free rule within a component of inflectional morphology that takes inputs from the rest of the morphological system and delivers its output to the syntactic component. This computational characterization may differ from many properties of rules informally called "regular" in traditional descriptive grammars. First, a regular rule in our sense need not be the only productive morphological phenomenon in a language; analogizing of irregular patterns (e.g., *brung*) can add new forms as well. Second, it need not apply to the majority of words in a given part-of-speech category; there is nothing in our characterization of regularity that directly pertains to the relative number of words involved, except that as a fully productive and general process, a regular rule is potentially applicable to an unlimited number of stems whereas irregular analogizing become less and less probable as one departs from a subclass prototype. As such, a regular rule may be the only available device for borrowings, neologisms, conversions, onomatopoeia, and other forms that cannot be analogized to existing irregulars (see Kim, Pinker, Prince, and Prasada, 1991; Prasada and Pinker, 1991). Thus a regular pattern is liable to assume majority status if such forms enter a language in large numbers, as has happened in English since the Middle English period (Pyles and Algeo, 1982). Third, a regular rule need not be totally insensitive to properties of the stem it affixes to: multiple rules may exist, each affixing an allomorph that applies to stems with a different property, though the conditions would be well-defined (e.g., ending in a vowel versus a consonant) and applicability to a stem within the class would be all or none, not proportional to the stem's global similarity to previously encountered stems. The most common hallmark of a regular process in our sense is *default status*, the ability to apply freely to any word not already linked to an irregular form, regardless of whether it is covered by memorization or analogy (e.g., *We rhumba'd all night*). The computational nature of the rule, as concatenation of an affix to a variable standing for the stem, renders it uniquely capable of playing this default role.

How Might a Regular Rule be Learned?

We have said nothing about how any particular regular rule, such as the English past tense rule, is acquired. Extraction of the suffixation pattern itself can be done by examining the phonological differences between forms like *walk* and *walked*, which can be accomplished by a variety of pattern extraction algorithms (see, e.g., MacWhinney, 1978; Pinker, 1984; Rumelhart & McClelland, 1986; Pinker & Prince, 1988). A greater challenge is to show how the child decides whether to internalize a pattern as a regular rule, or as a list of (possibly analogy-supporting) individual items such as *bend-bent*. How might this be done? Given the results of Chapter 11, children do not appear to depend on a regular pattern applying to a majority of tokens, nor to be influenced either by the relative or absolute number of types. Indeed children begin to overregularize with on the order of a hundred verbs in their vocabularies, and mark few or none of them for past tense before they begin to overregularize the irregulars, raising the possibility that the child does not need to process large amounts of input data to seize on the regular pattern. We lack evidence that would allow us to identify which cues children actually use to acquire a regular rule, but we can list several logical possibilities, each one a visible consequence of the regular nature of a rule and hence a potential cue to regularity (see Pinker and Prince, 1991, for further discussion). We can also check to see if any of these cues are clearly ruled out by existing evidence on the linguistic input available to children.

First, Pinker & Prince (1988) suggested that the crucial cue might be the ability of a morphological process to apply successfully to several kinds of stem, each belonging to a different competing irregular pattern. For example, *need-needed* exists despite *bleed-bled* and *feed-fed*; *blink-blinked* exists despite *drink-drank* and *sink-sank*; *seep-seeped* exists despite *sleep-slept* and *sweep-swept.* Indeed *all* the rhymes of irregular patterns in English (except perhaps for verbs ending in *ing*) can also be found in verb roots that are regular. This property is diagnostic of the fact that the regular rule does not carve out a set of phonological territories in the interstices of those claimed by irregular families, but can apply to anything at all. The phonological omnipotence of the regular rule is even more apparent in regular-irregular homophones, such as *lie-lay* (recline) versus *lie-lied* (fib), *fit-fit* (intransitive) versus *fit-fitted* (transitive), and *meet-met* versus *mete-meted.* (Of course, the homophones themselves are surely too rare to be the cues used by children.)

Second, the mere heterogeneity of the stem patterns that are heard to be regularly inflected (regardless of whether the inflection trumps some competing irregular pattern characteristic of such stems) may tell the child that the inflection is the product of a rule with either well-defined conditions or none at all. The Rumelhart-McClelland model is driven by a process roughly of this ilk, because the applicability of the regular suffix in the model depends on the variety of input patterns that are linked to the output nodes representing that suffix, and these links are strengthened with exposure to the relevant patterns. But the learning process we consider here is somewhat different. As mentioned, the range of generalization of the regular pattern in the Rumelhart-McClelland model is computed over the same representation that is used to represent the words to begin with. This representation has to have enough information to distinguish regulars from any irregular and for the full phonetic form of the past tense form to be reconstructed. This allows the potential conditions for the regular pattern to be sliced too finely; the regular pattern can be tied to highly specific combinations of stem patterns, with no across-the-board generalization to all non-irregular stems. If instead, the phonological content of words is represented in one, detailed kind of representation, but a small number of candidate conditions of a rule are represented in a separate, much sparser one, then a morphological pattern can be recognized as fully applicable as soon as it is seen to appear with every one of the stem conditions in this smaller, better-defined set. See Pinker (1984), Pinker & Prince (1988, 1991) and Prasada and Pinker (1991) for discussion.

Third, a very reliable cue for regularity is the ability of a process to apply to verbs that are derived from other categories such as nouns, adjectives, and names. These verbs do not have verb roots and hence have no base lexical entry to which an irregular memorized form can be attached; only a fully general regular rule can apply to them, acting as a default (see Pinker & Prince, 1988; Kim, Pinker, Prince, and Prasada, 1991; Kiparsky, 1982; Kim, Marcus, Hollander, and Pinker, 1991). For example *ring the city* (= "form a ring around") has the past tense form *ringed*, not *rang* (because this sense of *to ring* has a noun root, not a verb root that could be collapsed with the one underlying *ring-rang*). Hearing such forms could, in principle, tell a learner that the *-ed* affix is not stored with verb roots and hence must be a regular operation.

Fourth, even a single word with a highly unusual sound pattern, such as *out-Gorbachev'd* or *rhumba'd*, provides information about the generality and possible default status of a morphological pattern.

Fifth, the syntactic requirement that tense be marked *obligatorily* may impel the child's learning mechanism to seek a pattern that can provide the needed form under a wide range of circumstances. The fact that Adam, Eve, Sarah, and Abe were marking irregular verbs for tense a majority of the time near the period in which they were first controlling the regular suffix lends some credence to this suggestion. Of course obligatoriness itself cannot distinguish between regular rule and irregular storage-plus-analogy, but it might cause the child to promote the most general process to rule status.

Finally, aspects of the phonological properties of regular inflection might provide cues about its regularity. For example, the English past tense inflectional process consists of a suffixation, which is the same kind of process used for the third person singular inflection (*-s*), the progressive inflection (*ing*), the plural, and the possessive. This could signal to the learner that in English, suffixation is the process used quite generally for regular inflection, whereas the mutations seen in irregular verbs do not belong to any larger system (see Wurzel, 1989). Moreover a suffix that exerts no change upon a stem, or some change that is consistent across all stems, might be classified as regular, because it would suggest a process that treats the verb as an opaque variable rather than in terms of its phonological content (Kiparsky, 1982). In English these properties hold of regular inflection, and of course are confounded: adding a *-d* suffix leaves virtually all verbs unchanged (the exceptions are the irregular verbs *flee, say, hear, sell, tell*, and *do*).

As mentioned, this is a list of hypothetical cues, which we provide to allay possible suspicions that regularity in our sense is unlearnable, or learnable only from input sequences that our data already rule out. To decide which of these cues are in fact used by the child's learning mechanisms, it is necessary to determine which of them reliably accompany regular rules across different languages, which are in fact available in parental speech to children of the relevant ages, and which ones are controlled by children simultaneously with their treatment of the pattern as a regular process.

The actual set of regular verbs that children use to induce regular past tense inflection is, of course, difficult to determine from transcript data, which records a small fraction of the speech heard by children, and provides no information about what portion of that speech is actually used by the child. In Appendix we attempt to bracket the true list by providing one list that is likely to be too big, and another list that is likely to be too small. The maximal list contains all the regular verbs used at least once in the entire set of transcripts, in any tense, by Adam, Eve, Sarah, or the adults conversing with them. This list approximates the pool of verbs potentially occurring in past tense forms in parental speech to which the children might be attending before they deploy the regular inflection productively. We do not know, however, that each child heard each verb in this list, nor that each child heard each verb before acquiring the regular inflection, nor that each child had actually attended to the verbs they did hear. Thus we also provide a more conservative estimate, consisting only of the regular verbs that each child actually used (in any tense) in his or her transcripts preceding the first transcript with an overregularization.

The maximal list is consistent with several of the possibilities for cues to regularity entertained in this section. It contains many verbs that rhyme with irregulars (e.g., *ache*, *blind*, *blink*, *call*, *care*, *cheat*, *claw*, *cry*, *dare*, *die*, *end*, *exercise*, *fold*, *fry*, *guide*, *hatch*, *heat*, *invite*, *hand*, *kick*, *kid*, *land*, *lick*, *leak*, *like*, *live*, *mind*, *pee*, *peel*, *peep*, *pick*, *pin*, *play*, *pray*, *pretend*, *rake*, *reach*, *row*, *remind*, *repair*, *scare*, *scoot*, *sew*, *ski*, *smell*, *sneeze*, *snow*, *spell*, *spray*, *squeeze*, *stare*, *stay*, *steam*, *surprise*, *sway*, *tease*, *tick*, *tie*, *trick*, *trim*, *trust*, *turn*, *use*, *wet*, *yell*). Indeed, most of the irregular clusters of rhymes have regular counterparts in the list. In addition, there are a number of likely denominal verbs, including *bomb, bubble, chain, clip, color, comb, dust, end, fan, fish, glue, hammer, iron, lock, mail, paint, paste, pee, pump, rope, sail, screw, shovel, skate, ski, smoke, stamp, staple, steam, tape, thread, unbuckle, unbutton, unchain, unplug, unscrew, wee-wee*, and onomatopoeic verbs, which may have the same status (see Pinker and Prince, 1988): *bang, bark, bash, blast, bop, burp, clap, crack, crash, fizz, growl, howl, jabber, meow, pop, smash, swish, zip*, and *zoom*. Verbs with noncanonical and complex phonologies (especially, verbs that do not conform to canonical pattern for English basic words, namely monosyllables or polysyllables with initial stress) are represented by *allow, attach, belong, decorate,* *disappear, disturb, erase, excuse, fizz, meow, organize, prepare, pretend, recognize, remember, squoosh*, and *urinate*, among others.

Interestingly, even the minimal lists (regular verbs used by each child before overregularizing) exemplify most of these potential cues. The reader can confirm that each child controlled many verbs rhyming with a variety of irregulars. Adam and Sarah had acquired several denominal verbs (*fish, paint, plug, rain, rope, screw*, and *comb, paint, rain, smoke, wee-wee*, respectively), and onomatopoeic verbs (*crack, growl, knock, squeal, squeak*, and *knock* and *bump*, respectively). Adam and Sarah also had phonologically noncanonical verbs (*attach, exercise*, and *remember* for Adam; *organize* for Sarah). Even Eve, whose minimal list comes from only four transcripts, used 8 verbs rhyming with 6 kinds of irregulars *cry, kick, lick, pick, play, show, turn, use*, one likely deadjectival verb (*empty*), and an onomatopoeic form (*crack*). Thus the data suggest that each of the possible cues to regularity entertained in this section are available to children; whether they are used, necessary, or sufficient, however, remains unknown.

Relevance to Historical Change

Children regularize irregular verbs, especially the lower frequency ones, and the English language, over the centuries, has been regularizing irregular verbs, especially the lower frequency ones (Bybee, 1985). There must be some connection. Perhaps anticipating a trend, *Newsweek* (Gelman, 1986) attributed to Jill de Villiers the half-joke, "Leave children alone and they'd tidy up the English language."

But the remark can be turned around. Children, in fact, *are* left alone (Brown and Hanlon, 1970; Morgan and Travis, 1989; Marcus, 1991), but it does not really matter whether they are or not, for time is on their side. All they have to do is wait, and they will be adults in full custody of the language that is passed on to the next generation. If children really had a distaste for irregular verbs nothing could have stopped them from tidying up the English language long ago, yet we still have over a hundred irregular verbs, most quite secure, some of them transmitted as irregulars in a generation-to-generation chain of successful memorizations linking us to prehistoric peoples thousands of years in the past (Pyles and Algeo, 1982). Clearly children are not the relentless rule engines of earlier accounts, but are quite happy to learn irregulars, and why not? If one has the mental tools to acquire on the order of 50,000 words, each representing an arbitrary sound-meaning pairing, and the ability to link abstract features like tense to words with special markings, memorizing and retaining another few hundred words with the features built in should be no strain. And blocking their regularized counterparts is simply a consequence of the fact that stem and irregular are part of an organized grammatical system, with designated slots for each feature-marked variant of a word.

As Bybee and Slobin (1982b) point out, it is adults who bear most of the blame for the permanent regularization of irregulars in a language. They note that one cannot attribute historical change to children's overregularization errors unless children prefer these errors and continue to do so into adulthood, which for virtually all irregular verbs in a given generation, they do not. Given what we know about overregularization in children and adults (e.g., Ullman & Pinker, 1990, 1991), the following scenario seems more plausible. A weak irregular memory entry in adults can lead to occasional blocking failures, hence regularizations, for the same reason that children overregularize. Presumably this reduces the frequency of the irregular past tense form in the parent generation's speech further, and combined with an overall decline in all tense forms for the verb, it may erode to the point where one generation of children rarely hears it, and hence never cease to overregularize it, at which point it has changed to a regular.

Relevance to Connectionist Modeling of Language Acquisition

Although some of the questions treated in this monograph were first raised in Pinker and Prince's (1988) analysis of Rumelhart and McClelland's connectionist simulation of language acquisition, this monograph is not intended as an extension of their critique, and the points we make here are meant to shed light on the nature and causes of overregularization in general. They are not intended either to support or refute the entire set of connectionist models, nor the entire set of models based on grammatical rules, but pertain to many issues that are independent of that distinction, such as the relevant rarity of overregularization, its dependence on the frequency of irregular past forms, and its relation to tense marking in general.

In only three places did we explicitly test predictions of the Rumelhart-McClelland model. The first examined their suggestion that children's U-shaped development could be explained without a distinction between early rote storage of lexical items and later deployment of a productive morphological process. Rather, they suggested, a single pattern associator could display that transition given the assumption that irregulars are higher in frequency than regulars and that the child undergoes a vocabulary explosion causing the sudden acquisition of a large number of regulars and a concomitant change in the number of regular items submitted to the past tense learning device. Our results fail to find evidence either for a change in the input to the past tense system or for a vocabulary explosion at the right point in development that was assumed to cause it. In contrast, the data suggest that children use correct irregulars before overregularizing because they lack the productive mechanism generating regular forms altogether. These discoveries remove one of the more dramatic phenomena that had been adduced as evidence for the Rumelhart-McClelland model (e.g., Smolensky, 1988; Sampson, 1987). Of our other two tests, one supported the Rumelhart-McClelland model's prediction that irregular verbs are stored in a memory system that records patterns of mapping between stems and pasts and applies them in graded fashion to similar stems. The other test failed to support the model's prediction that the patterns shown by regular verbs are stored and generalized in this way.

A natural question arises: do the incorrect predictions of the Rumelhart-McClelland model apply to connectionist models of the acquisition of inflection in general, or only to their particular implementation? Theorists arguing for connectionism as a general solution to problems in cognitive science have, virtually unanimously, provided one answer: The problems are specific to the Rumelhart-McClelland model, and are primarily technological. The Rumelhart-McClelland model, an early effort in the recent revival of neural net modeling, was a feedforward network without a hidden layer; that is, each input node was connected directly to each output node, and activation passed from input to output in a single step. McClelland (1988), Plunkett and Marchman (1991, 1990), MacWhinney and Leinbach (in press), and Gasser and Lee (1990) have claimed that most of its problems are eliminated when higher-tech connectionist machinery is substituted. The usual suggested replacement is a feedforward network with one hidden layer of nodes separating input and output (e.g., Plunkett and Marchman, 1991; MacWhinney and Leinbach, in press), its connection weights modified by the error back-propagation learning algorithm (Rumelhart, Hinton, and Williams, 1986).

The claim that addition of a hidden layer is sufficient to remedy the Rumelhart-McClelland model's problems, especially its failed predictions about the dependence of overregularization on vocabulary balance, has never been demonstrated, however, and it is probably wrong. Egedi and Sproat (1991) have tested a PDP model that was trained on a sequence of English verbs similar to that fed to the Rumelhart-McClelland model, and they submitted their model to similar kinds of evaluation. The model enjoyed the advantages of a hidden layer whose weights were trained by the back-propagation algorithm, a more realistic phonological representation, and a more powerful mechanism converting the output node activations to pronounceable words. Nonetheless the behavior of the Egedi-Sproat model was virtually identical in the relevant respects to that of the Rumelhart-McClelland model. It displayed U-shaped development in its acquisition of irregulars at a point at which it was suddenly flooded with regulars, but showed no such trend with a constant input mixture. And it failed to produce a coherent output form for large numbers of new regular verbs on which it had been tested if those verbs differed phonologically from the common patterns in the training set.

Plunkett and Marchman (1991, 1990) claimed that their hidden-layer model, trained on a variety of schedules with artificial verbs, showed U-shaped development with a constant mixture of regular and irregular verbs as input. But they examined no psychological data and defined "U-shaped development" as any wiggle in a developmental curve, rather than the extended initial period of correct performance in children that psychologists had pointed to and that we confirmed quantitatively. Even this claim was abandoned in their subsequently-written paper (Plunkett and Marchman, 1990) which relied on a changing vocabulary balance as much as the Rumelhart-McClelland model did; moreover, in that paper they switched to a third and a fourth definition of U-shaped development (alternation between correct and irregular tokens, and immunity of early-learned irregulars from overregularization; see Section 10), again leaving the original phenomenon unexplained within their framework.

It is not surprising that the mere technological improvement of adding a hidden layer does not change the psychological fidelity of models with the basic design of the Rumelhart-McClelland model. The main point of the Pinker and Prince (1988) article was not that connectionist models are incapable of modeling psychological phenomena but that many of the key theoretical commitments in explaining such phenomena have little to do with whether a model is implemented in one or another kind of computational hardware but in more basic questions of representation and organization, such as: What is the format of representation for words? How many subsystems are used to map from stem to past? What kind of computation does each one do? What are its inputs and outputs? The problem with many connectionist models, they argued, was not inherent to connectionist machinery itself, but to the attempts of connectionist theorists to bypass these issues and attempt to use a single, all-purpose learning device for all linguistic and cognitive functions. In particular, Pinker and Prince questioned Rumelhart and McClelland's commitment (also adopted by most subsequent connectionist modelers) to a system that lacked distinct representations for words, that did not distinguish between regular rules and irregular storage but computed both within a single device, and that fed only phonological (and possibly semantic) information into the past tense system, ignoring morphological and syntactic structure. Obviously, adding a hidden layer of nodes to the Rumelhart-McClelland model has nothing to do with these issues.

The modeling efforts of MacWhinney and Leinbach (in press) demonstrate this point nicely. The authors do not attempt to model children's U-shaped development; they attribute it to a dissociation between rote storage of lexical entries and pattern extraction, as in the more traditional accounts of MacWhinney (1978), Pinker (1984), and Pinker and Prince (1988), and consider the lexical entries to belong to a separate system which they do not attempt to model. But they fail to acknowledge that the commitment to traditional linguistic distinctions exemplified in their model runs even deeper. MacWhinney and Leinbach's representation for the phonological content of the stem has a separate subset of hardware nodes for each position in a word, rather than representing order implicitly in a pattern of activation across a large set of context-sensitive nodes, as in the Rumelhart-McClelland model. More important, the architecture of their model, which they arrived at after a period of trial and error looking to see what worked, is clearly tailored to a qualitative architectural distinction between irregular storage and a stem-independent regular suffixation process. The main route of their model is a 4-layer (2-hidden layer) pattern associator with 200 nodes in each hidden layer, sufficient for memorizing the 180 irregular verbs in English (perhaps even by using some of the hidden nodes as "grandmother" cells, a direct implementation of the traditional notion of "rote memory"; see Scalettar and Zee, 1988). But they supplemented this conventional connectionist pattern associator with a distinct second route, which connected each input node directly with the output node coding the same phoneme and position, with an innate weight of 1.0 (the maximum), bypassing the intermediate layers and omitting all connections between one segment in the input and segments with different positions or different contents in the output. The bypass route is a second, innate, non-associative copying mechanism, tailor-made for regular verbs because the regular map simply copies the stem without internal modification and regardless of its content. Without calling attention to it, MacWhinney and Leinbach have designed a model with two distinct, innate pathways, one suited for irregular storage and analogy, the other suited for regular suffixation of the stem (see Prasada and Pinker, 1991, for further discussion).

In sum, we see no evidence that adding a hidden layer to a model with the basic design of the Rumelhart-McClelland model alleviates its empirical problems, in particular, the two that we examined in this monograph. At this point the models that leave its basic design untouched inherit its problems, and another model with better overall performance abandoned the radical assumption of a single associative map and adopted the traditional tenet of separate mechanisms for lexical entries, irregular links, and a nonassociative copying process for the regulars.

Although we have shown that existing connectionist models restricted to a single associative network do not account for the facts of children's level, onset, and lexical distribution of overregularization, we are not claiming that no connectionist model is capable of doing so. The set of possible connectionist models encompasses a wide variety of relative propensities for rote memory versus generalization, and includes models in which the balance between these tendencies changes during a training run. These tendencies are influenced by a variety of design parameters left open to the network creator, such as the coding scheme for the input features, their degree of probabilistic blurring, the topology of the network (e.g., the number of hidden layers, and the number of nodes in each one), the learning rate, the momentum factor, the temperature, the training schedule, and others. Thus conceivably someone might find a combination of design parameters that allow some unitary connectionist network to display childlike rote and regularization modes of behavior at different points in a realistic training sequence.

But even if such behavioral mimicry were achieved, this hypothetical model would still have to be evaluated against the full set of data relevant to the psychology of morphology (see Pinker, 1991; Pinker and Prince, 1991), and it is unlikely that a generic model tweaked with just the right parameter settings to display past tense overregularization would be consistent with the larger picture. For example, the English plural has very different vocabulary statistics from the past tense (only a few of the high frequency English nouns are irregular, so irregular forms cannot plausibly dominate early vocabulary), but it appears to develop with the same U-shaped sequence and low overregularization rate, and submits to the same grammatical treatment in terms of lexical storage and a default regular rule (Cazden, 1968; Pinker and Prince, 1988; Marcus, in preparation; Kim, Marcus, Hollander, and Pinker, 1991). Moreover there is a family of facts of a very different kind that independently favors a dual-mechanism (rule & rote) model over any unitary network, no matter how well it is designed to display rulelike and rotelike modes. These facts are summarized in the final section.

Evidence that Children Respect Qualitative Distinctions between Regular and Irregular Processes

Linguistic research has shown that regular and irregular inflected forms differ *qualitatively*, in terms of their sensitivity to qualitative grammatical distinctions and their relation to other grammatical processes. In at least three cases children have been shown to be sensitive to these distinctions in their patterns of regularization and overregularization.

First, Gordon (1986) noted that English compounds can contain irregular plurals (e.g., *mice-infested*) but not regular plurals (e.g., **rats-infested*). The phenomenon has been explained in terms of Kiparsky's (1982) Level-Ordering theory of morphology but the kernel of the explanation can be captured simply: the process forming such compounds takes as input stems stored in the lexicon, not complex forms created by an inflectional rule; if irregulars are stored in the lexicon, as we have suggested, they automatically can feed compounding. In an elicitation experiment, Gordon found that children produced novel compounds containing irregular plurals (*mice-eater*) but never novel compounds containing regular plurals (**rats-eater*). Nor did they form compounds out of their own overregularizations -- children who said *mouses* would nonetheless avoid saying **mouses-eater.*

Second, Stromswold (1990) looked at the three irregular verbs that are morphologically identical to auxiliaries: *do* (as in *do something* or *do it*), possessional *have*, and copula *be*. Each of these verbs is identical to an auxiliary not only in the stem form but in every single one of its irregularly inflected versions. *Do*, whether used as a main verb or as the auxiliary in negations, inversions, and emphatics, has the irregular forms *does, did*. The perfect auxiliary *have*, like possessional *have*, maps onto the irregular forms *has* and *had*. *Be*, whether serving as a copula verb, a progressive auxiliary, or a passive auxiliary, has the irregular forms *am, is, are, was, were, been*. Moreover, the semantic relations are the same in all cases: the relation between *have a book* and *had a book* is the same as the relation between *have eaten* and *had eaten*; *I am tired* is to *I was tired* as *I am resting* is to *I was resting*. Clearly there are too many of these parallelisms to be coincidental, and a parsimonious assumption is that the irregular forms of the main verb and of the auxiliary versions are stored in the same mechanism, because if they were not, at least one divergence among these 10 comparisons would be expected. But their susceptibility to overregularization is qualitatively different: in a sample of 40,000 child sentences containing these verbs, Stromswold found that the main verb versions are overregularized at rates comparable to those we have found here, whereas the auxiliary versions of the same verbs were *never* overregularized. This extreme conservatism was predicted, on the grounds of learnability considerations and other developmental data, by the theory in Pinker (1984) according to which the child's language acquisition mechanisms recognize auxiliary verbs to belong to a distinguished set of elements because of certain semantic and phonological properties, and as a result sequester their lexical representations from regular morphological operations. In other words, children appear to be capable of treating two verb forms identically in terms of irregular inflectional patterns, but qualitatively differently in terms of regular inflectional patterns.

Third, Kim, Marcus, Hollander, and Pinker (1991) tested children's sensitivity to the constraint mentioned in Section 14 that verbs derived from nouns and adjectives are regular even if homophonous to an irregular verb. For example, *to high-stick* (= "hit with a stick held high") has the past tense *high-sticked*, not *high-stuck*. Recall that the effect is a consequence of the fact that possessing an irregular past tense form is a property that holds of verb *roots*, not verbs in general (unless they have an irregular verb root as their head); verbs formed from nouns have a noun as their head, not an irregular verb root, hence there is no source for an irregular form and the default regular rule reliably applies. So in order to predict a verb's past tense form, its phonological properties do not suffice (nor do its semantic features; see Kim, Pinker, Prince, and Prasada, 1991); its morphological structure must be input to the relevant mechanisms. The effect has been demonstrated to be extremely robust in naive adults faced with inflecting novel denominal verbs (Kim, Pinker, Prince, and Prasada, 1991), and Kim, Marcus, Hollander, and Pinker (1991) showed it to hold in 3-8-year old children as well: children regularize denominal verbs homophonous with irregulars, such as *to fly* in the sense of "to cover a piece of paper with flies," more often than they overregularize the irregular verb itself, even when the irregular verb is used with nonstandard meanings and hence is equally unfamiliar.

These three phenomena suggest that it is necessary to attribute children's regularizations (including overregularizations) to a different mechanism than their irregulars, on the basis of the qualitatively different inputs and outputs the two patterns implicate: the production of irregulars but not overregularizations feeds compound formation; the production of irregulars but not overregularizations is fed by auxiliary verbs; the production of regularizations of verbs but not their irregular forms is fed by verbs derived from nouns. We have shown that exactly this kind of distinction between memory and rule mechanisms provides a simple account for a huge variety of facts about the rate, onset, and lexical properties of overregularization.

Appendices

:

Adam's Rates of Overregularization and Past Tense Marking

The columns correspond to: Age; Number of correct irregular past tense tokens; Number of stem+*ed* (e.g., *breaked*) overregularization tokens; Number of past+*ed* (e.g., *broked*) overregularization tokens; Rate of overregularization (proportion of irregular past tense form tokens that are overregularizations); Number of correct irregular past tense forms in obligatory past tense contexts; Number of stem forms of irregular verbs in obligatory past tense contexts; Irregular past marking rate (proportion of obligatory past tense contexts, not including overregularizations, where an irregular verb was supplied in the past tense); Number of correct regular past tense forms in obligatory past tense contexts; Number of stem forms of regular verbs in obligatory past tense contexts; Regular past marking rate (proportion of obligatory past tense contexts where a regular verb was supplied in the past tense). The last six columns are based on unpublished data provided by Courtney Cazden, described in Cazden (1966, 1968) and Brown (1973).

Age	**Corr Irreg**	**Overrg stem+*ed***	**Overrg past+*ed***	**Overrg Rate**	**Irreg Corr Oblig Cntexts**	**Irreg Stems Oblig Cntexts**	**Irreg Mark Rate**	**Reg Corr Oblig Cntexts**	**Reg Stems Oblig Cntexts**	**Reg Mark Rate**
2;3	40	0	0	0.00	15	15	0.50	0	5	0
2;4	14	0	0	0.00	4	11	0.27	4	24	0.14
2;5	51	0	0	0.00	38	12	0.76	1	27	0.04
2;6	90	0	0	0.00	54	8	0.87	1	16	0.06
2;7	58	0	0	0.00	60	11	0.85	0	11	0
2;8	61	0	0	0.00	44	16	0.73	4	13	0.24
2;9	31	0	0	0.00	26	21	0.55	0	12	0
2;10	36	0	0	0.00	31	3	0.91	0	5	0

2;11	39	2	0	0.05	31	6	0.84	9	11	0.45
3;0	74	0	0	0.00	47	18	0.72	16	11	0.59
3;1	115	0	0	0.00	64	2	0.97	8	0	1.00
3;2	103	0	0	0.00	51	10	0.84	5	7	0.42
3;3	181	0	0	0.00	108	10	0.92	51	10	0.84
3;4	132	3	0	0.02	80	6	0.93	15	7	0.68
3;5	124	3	0	0.02	95	1	0.99	14	2	0.88
3;6	108	1	0	0.01	60	0	1.00	14	0	1.00
3;7	93	1	1	0.02						
3;8	133	2	0	0.01						
3;9	85	3	0	0.03						
3;10	31	1	0	0.03						
3;11	76	1	0	0.01						
4;0	57	0	0	0.00						
4;1	43	1	0	0.02						
4;2	59	1	0	0.02						
4;3	42	1	0	0.02						
4;4	79	4	0	0.05						
4;5	47	1	0	0.02						
4;7	151	10	1	0.07						
4;8	40	2	0	0.05						
4;9	55	3	0	0.05						
4;10	68	1	2	0.04						
4;11	79	0	0	0.00						
5;2	51	3	0	0.06						

:

Eve's Rates of Overregularization and Past Tense Marking

The columns correspond to: Age; Number of correct irregular past tense tokens; Number of stem+*ed* (e.g., *breaked*) overregularization tokens; Number of past+*ed* (e.g., *broked*) overregularization tokens; Rate of overregularization (proportion of irregular past tense form tokens that are overregularizations); Number of correct irregular past tense forms in obligatory past tense contexts; Number of stem forms of irregular verbs in obligatory past tense contexts; Irregular past marking rate (proportion of obligatory past tense contexts, not including overregularizations, where an irregular verb was supplied in the past tense); Number of correct regular past tense forms in obligatory past tense contexts; Number of stem forms of regular verbs in obligatory past tense contexts; Regular past marking rate (proportion of obligatory past tense contexts where a regular verb was supplied in the past tense). The last six columns are based on unpublished data provided by Courtney Cazden, described in Cazden (1966, 1968) and Brown (1973).

Age	**Corr Irreg**	**Overrg stem+*ed***	**Overrg past+*ed***	**Overrg Rate**	**Irreg Corr Oblig Cntexts**	**Irreg Stems Oblig Cntexts**	**Irreg Mark Rate**	**Reg Corr Oblig Cntexts**	**Reg Stems Oblig Cntexts**	**Reg Mark Rate**
1;6	2	0	0	0	1	5	0.17	2	2	0.50
1;7	5	0	0	0	1	4	0.20	0	15	0.00
1;8	13	1	1	0.13	13	1	0.93	2	8	0.20
1;9	52	0	0	0.00	15	13	0.54	4	12	0.25
1;10	24	3	0	0.11	15	12	0.56	7	5	0.58
1;11	34	2	0	0.06	25	18	0.58	6	5	0.55
2;0	25	0	0	0.00	12	19	0.39	12	6	0.67
2;1	30	9	0	0.23	20	24	0.45	25	2	0.93
2;2	60	7	0	0.10	51	7	0.88	20	1	0.95
2;3	40	1	0	0.02	10	4	0.71	2	1	0.67

Sarah's Rates of Overregularization and Past Tense Marking

The columns correspond to: Age; Number of correct irregular past tense tokens; Number of stem+*ed* (e.g., *breaked*) overregularization tokens; Number of past+*ed* (e.g., *broked*) overregularization tokens; Rate of overregularization (proportion of irregular past tense form tokens that are overregularizations); Number of correct irregular past tense forms in obligatory past tense contexts; Number of stem forms of irregular verbs in obligatory past tense contexts; Irregular past marking rate (proportion of obligatory past tense contexts, not including overregularizations, where an irregular verb was supplied in the past tense); Number of correct regular past tense forms in obligatory past tense contexts; Number of stem forms of regular verbs in obligatory past tense contexts; Regular past marking rate (proportion of obligatory past tense contexts where a regular verb was supplied in the past tense). The last six columns are based on unpublished data provided by Courtney Cazden, described in Cazden (1966, 1968) and Brown (1973).

Age	**Corr Irreg**	**Overrg stem+*ed***	**Overrg past+*ed***	**Overrg Rate**	**Irreg Corr Oblig Cntexts**	**Irreg Stems Oblig Cntexts**	**Irreg Mark Rate**	**Reg Corr Oblig Cntexts**	**Reg Stems Oblig Cntexts**	**Reg Mark Rate**
2;3	8	0	0	0	1.5	0	1	0	0.5	0
2;4	19	0	0	0	4.5	8	0.36	0	2	0
2;5	21	0	0	0	8.5	3	0.74	1	0	1.00
2;6	26	0	0	0	16	14	0.53	2	1.5	0.57
2;7	57	0	0	0	20.5	13	0.61	0	4.5	0
2;8	20	0	0	0	10	2	0.83	0	0.5	0
2;9	80	0	0	0	24	6	0.80	9	6	0.60
2;10	25	1	0	0.04	7	3	0.70	10	4	0.71
2;11	27	0	0	0	10	3.5	0.74	1.5	0.5	0.75
3;0	46	0	0	0	12	5.5	0.69	3.5	1	0.78

3;1	34	0	0	0	35	0	1	4	2.5	0.62
3;2	52	0	0	0	24	2	0.92	3.5	2	0.64
3;3	39	1	0	0.02	24.5	2	0.92	2.5	2	0.56
3;4	20	0	0	0	19.5	1	0.95	1	1	0.5
3;5	44	1	0	0.02	17	6	0.74	3	0	1.00
3;6	35	1	0	0.03	16.5	4.5	0.79	2.5	1	0.71
3;7	61	0	0	0	34	4	0.89	5.5	0.5	0.92
3;8	63	1	0	0.02	38.5	4.5	0.90	8	2.5	0.76
3;9	27	0	0	0	23.5	0	1	2.5	0	1.00
3;10	64	3	0	0.04	45.5	1	0.98	21.5	1	0.96
3;11	67	2	0	0.03	42.5	5.5	0.89	6	0	1.00
4;0	30	0	1	0.03	21.5	4.5	0.83	4	1	0.80
4;1	67	1	0	0.01	45	1	0.98	13	1	0.93
4;2	122	9	1	0.08	50	4	0.93	22	0	1.00
4;3	41	1	0	0.02						
4;4	76	1	0	0.01						
4;5	54	2	0	0.04						
4;6	70	0	0	0						
4;7	71	5	1	0.08						
4;8	64	12	0	0.16						
4;9	59	7	1	0.12						
4;10	56	4	0	0.07						
4;11	78	4	0	0.05						
5;0	58	2	0	0.03						
5;1	36	3	0	0.08						

Abe's Rates of Overregularization and Past Tense Marking

The columns correspond to: Age; Number of correct irregular past tense tokens; Number of stem+*ed* (e.g., *breaked*) overregularization tokens; Number of past+*ed* (e.g., *broked*) overregularization tokens; Rate of overregularization (proportion of irregular past tense form tokens that are overregularizations); Irregular past marking rate (correct irregular pasts as a proportion of the sum of correct irregular pasts and stem forms in obligatory past tense contexts); Regular past marking rate (correct regular pasts as a proportion of the sum of correct regular pasts and stem forms in obligatory contexts.) The last two columns are taken from the Appendices to Kuczaj (1976; see also Kuczaj, 1977a).

Age	**Corr Irreg**	**Overrg stem+*ed***	**Overrg past+*ed***	**Overrg Rate**	**Irreg Mark Rate**	**Reg Mark Rate**
2;5	9	2	2	0.31	.31	.76
2;6	24	1	0	0.04	.58	.63
2;7	27	16	2	0.40	.69	.87
2;8	52	12	1	0.20	.81	.80
2;9	102	36	6	0.29	1.00	.98
2;10	44	34	6	0.48	1	1
2;11	76	31	11	0.36	.96	1
3;0	24	16	2	0.43	1	1
3;1	64	19	11	0.32	.99	1
3;2	82	8	5	0.14	.99	1
3;3	91	36	9	0.33	.97	1
3;4	107	38	4	0.28	1	1
3;5	89	31	1	0.26	1	1

3;6	74	12	1	0.15	1	1
3;7	58	4	0	0.06	1	1
3;8	88	16	2	0.17	1	1
3;9	80	22	2	0.23	1	1
3;10	69	14	2	0.19	1	1
3;11	75	21	3	0.24	1	1
4;0	13	2	6	0.38	1	1
4;1	67	19	7	0.28	1	1
4;2	42	7	4	0.21	1	1
4;3	39	14	10	0.38	1	1
4;4	16	10	0	0.38	1	1
4;5	29	3	0	0.09	1	1
4;6	68	3	0	0.04	1	1
4;7	59	3	1	0.06	1	1
4;8	38	0	0	0	1	1
4;9	26	0	0	0	1	1
4;10	62	16	0	0.21	1	1
4;11	47	6	1	0.13	1	1
5;0	45	13	0	0.22	1	1

:
Adam's Irregular Verbs

Verb	**Correct**	**Stem+*ed***	**Past+*ed***	**Overreg Rate**
be	137	0	0	0
beat	1	2	0	0.67
bend	1	0	0	0
bite	12	0	0	0
bleed	0	0	0	--
blow	0	1	0	1
break	97	0	0	0
bring	17	0	0	0
build	0	0	0	--
buy	10	0	0	0
catch	73	2	0	0.03
choose	0	0	0	--
come	51	1	0	0.01
cut	11	1	0	0.08
dig	0	0	0	--
do	93	1	0	0.01
draw	1	4	0	0.80
drink	0	0	0	--
drive	1	1	0	0.50
eat	21	0	0	0
fall	182	2	0	0.02
feed	0	0	0	--

feel	0	3	0	1
fight	0	0	0	--
find	60	0	0	0
fly	1	0	0	0
forget	27	0	0	0
freeze	0	0	0	--
get	586	0	0	0
give	19	0	0	0
go	125	0	0	0
grind	0	0	0	--
grow	1	2	0	0.67
hang	0	0	0	--
have	56	0	0	0
hear	53	0	0	0
hide	0	0	0	--
hit	35	0	0	0
hold	0	1	0	1
hurt	28	0	0	0
keep	0	0	0	--
know	1	0	0	0
leave	15	0	0	0
lend	0	0	0	--
let	3	0	0	0
lie	0	0	0	--
light	0	0	0	--
lose	50	0	0	0

make	141	7	0	0.05
mean	0	0	0	--
meet	1	0	0	0
put	69	0	0	0
read*				
rend	0	0	0	--
ride	5	0	0	0
ring	0	0	0	--
rise	0	0	0	--
run	9	0	1	0.10
say	129	0	0	0
see	104	0	0	0
send	0	0	0	--
shake	0	0	0	--
shoot	35	0	0	0
shrink	0	0	0	--
shut	0	0	0	--
sing	0	0	0	--
sit	1	0	0	0
sleep	0	1	0	1
slide	0	0	0	--
spend	0	0	0	--
spin	0	0	0	--
spit	0	0	0	--
stand	0	1	0	1
steal	4	0	2	0.33

stick	2	4	0	0.67
strike	0	0	0	--
string	0	0	0	--
sweep	0	0	0	--
swim	0	0	0	--
swing	0	0	0	--
take	62	1	1	0.03
teach	0	1	0	1
tear	1	0	0	0
tell	59	0	0	0
think	33	0	0	0
throw	3	3	0	0.50
wake	1	1	0	0.50
wear	0	0	0	--
win	0	4	0	1
wind	0	0	0	--
write	17	0	0	0

* Analyzed only for Abe.

:

Eve's Irregular Verbs

Verb	**Correct**	**Stem+*ed***	**Past+*ed***	**Overreg Rate**
be	22	0	0	0
beat	0	0	0	--
bend	0	0	0	--
bite	7	0	0	0
bleed	0	0	0	--
blow	0	0	0	--
break	15	0	0	0
bring	2	0	0	0
build	0	0	0	--
buy	3	0	0	0
catch	6	0	0	0
choose	0	0	0	--
come	23	4	0	0.15
cut	0	0	0	--
dig	0	0	0	--
do	7	2	0	0.22
draw	0	0	0	--
drink	0	1	0	1
drive	0	0	0	--
eat	0	0	0	--
fall	2	8	0	0.80
feed	0	0	0	--

feel	0	0	0	--
fight	0	0	0	--
find	1	0	0	0
fly	0	0	0	--
forget	18	0	0	0
freeze	0	0	0	--
get	55	0	0	0
give	3	0	0	0
go	20	5	0	0.20
grind	0	0	0	--
grow	0	0	0	--
hang	0	0	0	--
have	20	0	0	0
hear	0	0	0	--
hide	0	0	0	--
hit	1	0	0	0
hold	0	0	0	--
hurt	5	0	0	0
keep	0	0	0	--
know	0	0	0	--
leave	1	0	0	0
lend	0	0	0	--
let	0	0	0	--
lie	0	0	0	--
light	0	0	0	--
lose	3	0	0	0

make	7	0	0	0
mean	0	0	0	--
meet	0	0	0	--
put	38	0	0	0
read*				
rend	0	0	0	--
ride	0	0	0	--
ring	0	0	0	--
rise	0	0	0	--
run	0	0	0	--
say	10	0	0	0
see	3	1	0	0.25
send	1	0	0	0
shake	0	0	0	--
shoot	0	0	0	--
shrink	0	0	0	--
shut	1	0	0	0
sing	0	0	0	--
sit	4	0	0	0
sleep	0	0	0	--
slide	0	0	0	--
spend	0	0	0	--
spin	0	0	0	--
spit	0	0	0	--
stand	0	0	0	--
steal	0	0	0	--

stick	1	0	0	0
strike	0	0	0	--
string	0	0	0	--
sweep	0	0	0	--
swim	0	0	0	--
swing	0	0	0	--
take	0	0	0	--
teach	0	0	0	--
tear	0	0	1	1
tell	2	0	0	0
think	1	0	0	0
throw	0	1	0	1
wake	0	0	0	--
wear	0	1	0	1
win	0	0	0	--
wind	0	0	0	--
write	1	0	0	0

* Analyzed only for Abe.

:
Sarah's Irregular Verbs

Verb	**Correct**	**Stem+*ed***	**Past+*ed***	**Overreg Rate**
be	182	0	0	0
beat	0	0	0	--
bend	0	1	0	1
bite	11	0	0	0
bleed	0	0	0	--
blow	1	0	0	0
break	96	0	0	0
bring	9	0	0	0
build	0	0	0	--
buy	18	3	0	0.14
catch	31	1	1	0.06
choose	0	0	0	--
come	15	3	1	0.21
cut	6	0	0	0
dig	0	0	0	--
do	127	0	0	0
draw	0	0	0	--
drink	4	0	0	0
drive	0	0	0	--
eat	14	0	0	0
fall	10	1	0	0.09
feed	0	0	0	--

feel	0	0	0	--
fight	0	2	0	1
find	27	0	0	0
fly	2	1	2	0.6
forget	30	0	0	0
freeze	1	0	0	0
get	434	0	0	0
give	33	0	0	0
go	66	6	0	0.08
grind	0	0	0	--
grow	1	1	0	0.50
hang	0	1	0	1
have	137	0	0	0
hear	10	5	0	0.29
hide	2	1	0	0.33
hit	27	0	0	0
hold	0	0	0	--
hurt	9	1	0	0.10
keep	1	0	0	0
know	5	0	0	0
leave	8	0	0	0
lend	1	0	0	0
let	1	0	0	0
lie	0	0	0	--
light	2	0	0	0
lose	19	1	0	0.05

make	52	8	0	0.13
mean	1	0	0	0
meet	2	0	0	0
put	55	1	0	0.02
read*				
rend	0	0	0	--
ride	0	0	0	--
ring	0	0	0	--
rise	0	0	0	--
run	4	2	0	0.33
say	101	0	0	0
see	59	0	0	0
send	0	0	0	--
shake	2	0	0	0
shoot	0	0	0	--
shrink	1	0	0	0
shut	3	0	0	0
sing	1	0	0	0
sit	0	1	0	1
sleep	2	0	0	0
slide	0	3	0	1
spend	0	0	0	--
spin	2	0	0	0
spit	0	0	0	--
stand	0	0	0	--
steal	0	2	0	1

stick	1	0	0	0
strike	3	0	0	0
string	0	0	0	--
sweep	0	0	0	--
swim	0	1	0	1
swing	0	0	0	--
take	35	2	0	0.05
teach	5	0	0	0
tear	0	0	0	--
tell	25	1	0	0.04
think	10	0	0	0
throw	3	7	0	0.70
wake	3	2	0	0.40
wear	1	0	0	0
win	4	3	0	0.43
wind	1	0	0	0
write	1	0	0	0

* Analyzed only for Abe.

:

Abe's Irregular Verbs

Verb	Correct	Stem+*ed*	Past+*ed*	Overreg Rate
be*				
beat	0	0	0	--
bend	1	1	0	0.50
bite	3	4	0	0.57
bleed	0	0	0	--
blow	4	9	0	0.69
break	38	4	12	0.30
bring	2	5	1	0.75
build	1	8	0	0.89
buy	7	5	0	0.42
catch	22	5	1	0.21
choose	1	1	0	0.50
come	20	52	4	0.74
cut	15	12	0	0.44
dig	2	3	0	0.60
do*				
draw	1	6	0	0.86
drink	5	5	0	0.50
drive	0	2	0	1
eat	82	18	2	0.20
fall	72	54	3	0.44
feed	0	1	0	1

feel	5	11	0	0.69
fight	1	2	0	0.67
find	143	1	2	0.02
fly	5	4	0	0.44
forget	67	0	0	0
freeze	1	0	0	0
get	194	4	50	0.22
give	5	2	0	0.29
go	117	60	4	0.35
grind	0	0	0	--
grow	2	6	0	0.75
hang	0	4	0	1
have*				
hear	13	27	0	0.68
hide	2	0	0	0
hit	16	8	0	0.33
hold	0	4	0	1
hurt	16	2	0	0.11
keep	1	0	0	0
know	11	6	0	0.35
leave	13	2	0	0.13
lend	0	0	0	--
let	0	0	0	--
lie	0	0	0	--
light	0	0	0	--
lose	8	1	0	0.11

make	150	24	3	0.15
mean	3	1	0	0.25
meet	0	0	0	--
put	77	11	0	0.12
read	1	1	0	0.50
rend	0	0	0	--
ride	2	1	0	0.33
ring	0	0	0	--
rise	0	0	0	--
run	4	6	0	0.60
say	282	3	0	0.01
see	166	4	3	0.04
send	8	1	0	0.11
shake	2	0	0	0
shoot	10	2	1	0.23
shrink	0	0	0	--
shut	0	0	0	--
sing	2	1	0	0.33
sit	1	1	0	0.50
sleep	1	2	0	0.67
slide	0	0	0	--
spend	0	1	0	1
spin	0	0	0	--
spit	0	2	1	1
stand	1	2	0	0.67
steal	3	0	0	0

stick	1	1	0	0.50
strike	0	0	0	--
string	0	1	0	1
sweep	0	2	0	1
swim	0	2	0	1
swing	0	3	0	1
take	21	6	3	0.30
teach	3	0	0	0
tear	0	0	0	--
tell	52	7	0	0.12
think	75	9	9	0.19
throw	5	12	0	0.71
wake	3	1	0	0.25
wear	0	3	0	1
win	16	9	0	0.36
wind	0	0	0	--
write	1	7	0	0.88

*Analyzed only for Adam, Eve, and Sarah.

:

Other Children's Overregularization Rates

Verb	ANC	Allsn	April	BOM	BRD	CHJ	DED	GAT	JOB	JUB	KIF	MAA	MIM	Naomi	Nat	Nthnl	Peter	TOS	TRH	VOH	ZOR
be*																					
beat*																					
bend	--	--	--	--	--	--	--	0	--	--	--	--	0	0.33	--	--	--	--	--	--	--
bite	--	1	0	0	0	0	0	0	0	0.12	0	0	0	0	--	0	0	0	0	--	--
bleed	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
blow	0	--	--	--	0	--	0	0	--	--	--	--	0	--	--	0	0	--	--	--	--
break	0	0	0	0	0	0	0	0	0	--	--	0	--	0	0	0.29	0.075	--	0	0	0
bring	--	--	--	--	--	--	0	0	0	--	--	0	0	0	0	--	0	--	--	--	--
build	0	--	--	--	--	0	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

buy	--	--	1	--	--	0	0	0	0	0	0	--	0	0	0	0	0	0	--	--	0
catch	--	--	--	1	0	0.67	0	0	--	--	--	1	--	0	0	0.25	0	0	--	--	--
choose	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
come	0	0	0.67	0	0	0	1	0	0	0	--	0	0	0	0	0	0	0	0	0	0
cut*																					
dig	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
do*																					
draw	1	--	--	0	--	--	--	1	--	--	--	0	--	1	0	--	0	--	1	--	--
drink	--	--	--	0	--	--	--	--	0	--	--	--	--	0	--	1	1	0	--	--	--
drive	--	--	1	--	--	--	--	--	--	--	--	--	--	--	--	0	0	--	--	--	--
eat	0	0	0	0	0	0	0	0	0	0	0	--	--	0	--	0.25	0	0	0	0	0
fall	1	0	0.43	0	0	0	--	0	0	--	0	0	--	0.04	0	0	0.10	0	1	--	--
feed	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0	--	--
feel	--	--	--	0	--	0	--	0	--	--	--	--	--	--	--	--	--	0	--	--	--
fight	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.5	--	--
find	0	--	0	0	0	0	0	0	0	0	0	0	--	0.2	--	0	0	--	--	--	0
fly	--	--	--	--	--	0	--	1	--	--	--	0	--	0.75	--	--	--	--	0	--	--
forget	0	--	0	0	0	0	0	0	0	--	0	--	0	0	--	0	0	0	--	--	0
freeze	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0	--	--
get	0	0	0	0	0	0	0	0	0	0	0	0	0	0.01	0	0	0	0	--	0	0
give	--	0	0	0	--	0	0	0	0	0	0	0	0	0	0	0	0.12	0	0	--	0
go	0	0	0	--	--	--	--	0	0	0	--	0	--	0.05	0	0.15	0.07	0	--	0	0
grind	--	--	--	0	0	--	--	0	0	0	--	--	--	--	--	--	--	--	--	--	0
grow	--	--	--	--	0.67	--	--	--	--	0	--	0	--	--	--	--	--	--	--	1	--
hang	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
have*																					
hear	0	--	--	0	0	0	0	0	0	0	0	0	0	0	--	--	0	0	--	0	--
hide	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0	--	--
hit*																					
hold	--	--	--	0	--	--	--	0	--	--	--	--	--	1	--	--	--	--	--	--	--
hurt*																					
keep	--	--	--	--	--	--	0	1	--	--	--	--	--	--	--	--	--	--	--	--	--
know	0	--	--	0	--	--	0	--	--	0	--	--	0	0	--	--	--	--	0	--	--
leave	--	0	--	0	0	0	0	0	--	--	--	--	0	--	--	0	0	0	--	--	0
lend	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0	--	--
let*																					
lie	--	--	--	--	0	0	--	0	--	--	--	--	--	--	--	--	--	--	--	--	--
light	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
lose	--	0	--	0	0	--	--	0	0	--	--	0	0	0	0	0	0	0	--	--	0
make	--	0	0	0	0	0	0	0.1	0	0	0	0	0	0.17	--	0	0	0	0	--	0
mean	0	--	--	--	--	0	--	--	0	--	--	--	--	--	--	--	--	--	0	--	--
meet	--	--	--	--	--	--	--	--	--	0	0	--	--	--	--	--	--	--	--	--	--
put*																					
read*																					
rend	--	--	--	--	--	--	0	--	--	--	--	--	--	--	--	--	--	--	--	--	--
ride	--	--	--	--	--	--	--	0	--	1	--	--	--	1	--	0	0	--	--	--	--
ring	--	--	--	--	--	--	--	1	--	--	--	--	--	--	--	--	--	--	--	--	--
rise	--	--	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--	--	--	--	--
run	--	--	--	--	--	0.4	0.5	--	--	0	--	--	--	0	--	0	--	--	--	0	0

say	0	0	0	0	--	0	0	0	0	0	0	0	--	0	--	0	0	0	--	0	0
see	0	--	0	0	--	0	0	0	0	0	0	0	--	0.29	0	0	0	0	--	0	--
send	--	--	--	0	--	--	1	--	0	0	--	--	--	--	--	--	--	--	--	--	--
shake	--	--	--	--	--	--	--	1	--	--	--	--	--	--	--	--	--	--	--	--	--
shoot	--	--	--	--	--	--	--	0	--	1	0	--	--	0	--	--	--	--	--	--	0
shrink	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--	--	--	--	--	--	--
shut*																					
sing	--	--	--	--	--	--	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--
sit	--	--	--	--	--	--	--	--	--	--	0	--	--	0	--	0	0	--	--	--	--
sleep	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0	--	--	--	--	--
slide	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
spend	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
spin	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
spit	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
stand	--	--	--	--	--	--	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--
steal	--	--	--	--	--	--	--	--	--	1	--	--	--	--	--	1	--	0	--	--	--
stick	0	0	--	--	--	--	--	--	--	--	--	0.25	--	0.08	--	0	0	0	--	--	0
strike	--	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--	--	--	--	--	--
string	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
sweep	--	--	--	--	--	--	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--
swim	--	--	0	--	--	--	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--
swing	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
take	--	--	0	--	--	--	--	--	0	0	--	--	--	0	0	0	0.13	0	--	0	0
teach	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--	--	0	--	--	0	--
tear	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
tell	0	--	--	--	--	--	--	--	0	0	--	0	--	0	--	0	0.33	0	--	0	0
think	0	--	--	--	--	--	--	0	--	0	--	0	--	0	--	--	0	0	--	0	0
throw	--	1	--	--	--	--	--	0	--	0	--	0	--	0.44	--	0.50	0	--	--	--	--
wake	--	--	--	--	--	--	--	--	--	--	--	0	--	0	--	--	0.60	--	--	--	--
wear	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
win	--	--	--	--	--	--	--	--	0	--	--	--	--	--	0	--	0	--	--	0	0
wind	--	--	--	--	--	--	--	0	--	--	--	--	--	--	--	--	--	--	--	--	--
write	--	--	--	--	--	--	--	--	0	1	--	--	--	1	--	--	0	--	--	--	0

* Analyzed only for Adam, Eve, Sarah, and/or Abe.

:

Aggregate Measures of Verbs' Overregularization Rates and Frequencies in Parental Speech

Verb	Mean Standardized Overregularization Rate (z-score)	Mean z Converted Back to Overregularization Rate	Parental Frequency
be	-0.48	0.03	5.81
beat	--	--	0.70
bend	--	--	0.34
bite	-0.55	0.01	3.10
bleed	--	--	0
blow	1.66	0.40	1.20
break	0.20	0.14	2.32
bring	-0.43	0.03	2.11
build	--	--	0.29
buy	0.16	0.14	2.19
catch	0.04	0.12	1.58
choose	--	--	0
come	0.35	0.17	2.88
cut	1.81	0.42	4.09
dig	--	--	0
do	-0.25	0.07	6.39
draw	--	--	0.75
drink	0.80	0.25	1.10
drive	--	--	1.34
eat	-0.50	0.02	2.24
fall	0.70	0.23	2.65

feed	--	--	0
feel	1.64	0.39	0.61
fight	--	--	1.25
find	-0.20	0.07	2.25
fly	--	--	0.62
forget	-0.60	0.01	2.30
freeze	--	--	0
get	-0.19	0.08	4.80
give	-0.47	0.03	2.70
go	0.27	0.16	3.69
grind	--	--	0
grow	--	--	0
hang	--	--	1.39
have	-0.48	0.03	5.10
hear	0.66	0.22	2.15
hide	--	--	0.69
hit	-0.30	0.06	3.93
hold	--	--	0.59
hurt	-0.40	0.04	4.07
keep	--	--	1.22
know	0.14	0.13	2.14
leave	-0.63	0.00	2.75
lend	--	--	0
let	--	--	5.19
lie	--	--	0.98
light	--	--	0

lose	-0.40	0.04	2.39
make	0.30	0.16	3.26
mean	--	--	1.45
meet	--	--	1.39
put	-0.55	0.01	6.27
read	--	--	4.08
rend	--	--	0
ride	--	--	0.45
ring	--	--	0
rise	--	--	0
run	2.51	0.54	2.08
say	-0.31	0.06	3.95
see	-0.56	0.01	3.11
send	--	--	0.92
shake	--	--	0.92
shoot	-0.42	0.04	1.28
shrink	--	--	0
shut	--	--	3.02
sing	--	--	0.51
sit	--	--	1.58
sleep	--	--	0.92
slide	--	--	0
spend	--	--	1.10
spin	--	--	0
spit	--	--	0
stand	--	--	-1.10

steal	--	--	0.59
stick	-0.06	0.10	2.32
strike	--	--	0
string	--	--	0
sweep	--	--	-0.69
swim	--	--	-1.39
swing	--	--	0
take	0.30	0.16	3.08
teach	--	--	1.16
tear	--	--	0.92
tell	-0.58	0.01	2.97
think	-0.51	0.02	3.60
throw	3.06	0.64	1.44
wake	--	--	1.10
wear	--	--	0.85
win	0.17	0.14	1.64
wind	--	--	0
write	-0.43	0.03	1.23

Regular Verbs Available to Children

Regular Verbs Used at Least Once by Adam, Eve, Sarah, or the Adults Conversing With Them

ache, act, add, aim, allow, answer, ask, attach, back, bake, balance, bang, bark, bash, believe, belong, blast, bless, blind, blink, bob, boil, bomb, bop, borrow, bother, bounce, bow, bowl, box, breathe, brush, bubble, bump, burn, burp, bury, bust, call, camp, care, carry, carve, chain, change, chase, cheat, check, chew, chirp, choke, chop, clap, claw, clean, climb, clip, close, color, comb, cook, cool, copy, cough, count, cover, crack, crash, crawl, criss-cross, cross, crush, cry, curl, dance, dare, dash, decorate, dial, die, dip, dirty, disappear, disturb, dodge, drag, dress, dribble, drill, drip, drool, drop, drown, dry, dump, dust, empty, end, erase, excuse, exercise, faint, fan, fasten, fetch, file, fill, finish, fish, fix, fizz, flash, flip, float, fold, follow, fool, frighten, fry, glue, grab, growl, guess, guide, hammer, hand, handle, happen, hatch, hate, heat, help, holler, hook, hop, hope, howl, hug, hunt, hurry, hush, invite, iron, itch, jabber, joke, juggle, jump, kick, kid, kill, kiss, knock, land, last, laugh, leak, lean, learn, lick, lift, like, listen, live, load, lock, look, love, mail, manage, march, marry, mash, match, matter, measure, melt, meow, mess, mind, miss, mix, mock, move, murder, name, nap, need, obey, open, organize, own, pack, paint, pardon, park, pass, paste, pat, pay, pee, pee-pee, peek, peel, peep, perch, pick, pin, pinch, place, plant, play, plug, point, poke, polish, pop, pour, pout, practice, pray, prepare, press, pretend, print, promise, pull, pump, punch, punish, push, race, rain, raise, rake, reach, reattach, recognize, remember, remind, repair, rest, rinse, rip, rock, roll, rope, row, rub, ruin, rustle, sail, salute, save, scare, scoop, scoot, score, scratch, scream, screw, scribble, seem, serve, sew, share, sharp, sharpen, shave, shop, shout, shove, shovel, show, shrug, sigh, sip, skate, ski, skip, slap, slip, smack, smart, smash, smell, smile, smoke, smooth, snap, sneeze, sniff, snow, snuggle, soak, sound, spank, spell, spill, splash, spoil, spray, squash, squeak, squeal, squeak, squeeze, squirt, squish, squoosh, stab, stamp, staple, stare, start, starve, stay, steam, steer, step, stir, stop, straighten, stretch, study, suck, surprise, swallow, sway, swish, switch, talk, tangle, tape, taste, tease, thank, thread, tick, tickle, tie, tighten, tip, touch, trap, trick, trim, trip, trust, try, tuck, turn, twist, type, unbuckle, unbutton, unchain, unplug, unscrew, untangle, untie, urinate, use, visit, wag, wait, walk, want, wash, watch, wave, wee-wee, weigh, wet, whirl, whisper, whistle, wiggle, wink, wipe, wish, wobble, wonder, work, worry, wrap, wreck, yawn, yell, zip, zoom

Regular Verbs Used by Adam in the Transcripts Preceding His First Overregularization

ask, attach, bake, belong, bless, bow, burn, call, care, carry, change, check, chew, climb, close, cook, copy, count, crack, crawl, cross, cry, dance, drip, drop, dry, dump, excuse, exercise, fasten, finish, fish, fix, fold, frighten, growl, hand, happen, help, hop, hug, hurry, jabber, jump, kick, kiss, knock, laugh, like, listen, live, look, love, march, mess, miss, mix, move, need, open, pack, paint, park, pay, peek, pick, pinch, play, plug, point, pour, press, pretend, pull, push, rain, reach, remember, rock, roll, rope, save, screw, shine, shop, show, skip, slip, smoke, snow, spank, spill, squeal, squeak, squeeze, stay, step, stir, stop, talk, taste, thank, tickle, tie, touch, try, turn, use, wait, walk, want, wash, watch, whisper, wipe, wonder, work

Regular Verbs Used by Eve in the Transcripts Preceding Her First Overregularization

answer, bang, change, climb, crack, cry, dance, drop, empty, fix, happen, help, jump, kick, lick, like, look, move, open, pardon, pick, play, show, spill, step, taste, thank, touch, turn, use, wait, want, watch, wipe,

Regular Verbs Used by Sarah in the Transcripts Preceding Her First Overregularization

bounce, bump, call, clean, close, comb, count, cover, cross, cry, dance, die, dress, dump, fix, happen, help, hug, jump, knock, laugh, lift, like, look, love, need, open, organize, paint, park, pat, peek, pick, pinch, play, poke, pour, pray, press, pull, push, rain, rip, rock, roll, shop, show, smoke, spank, stay, step, stop, talk, taste, tease, thank, tie, touch, turn, use, wait, walk, want, wash, watch, wee-wee, work, yawn

References

Adams, S. (1938) Analysis of verb forms in the speech of young children, and their relation to the language learning process. *Journal of Experimental Education, 7*, 141-144.

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

Anisfeld, M. & Gordon, M. (1968) On the psychophonological structure of English inflectional rules. *Journal of Verbal Learning and Verbal Behavior, 7*, 973-979.

Anisfeld, M. & Tucker, G. R. (1967) The English pluralization rules of six-year old children. *Child Development, 38*, 1201-1217.

Anisfeld, M., Barlow, J., & Frail, C. M. (1968) Distinctive features in the pluralization rules of English speakers. *Language and Speech, 11*, 31-37.

Aronoff, M. (1976) *Word formation in generative grammar.* Cambridge, Mass.: MIT Press.

Bateman, W. G. (1916) The language status of three children at the same ages. *Pedagogical Seminary, 23*, 211-240.

Berko, J. (1958) The child's learning of English morphology. *Word, 14*, 150-177.

Bever, T. G. (1975) Psychologically real grammar emerges because of its role in language acquisition. In D. Dato (Ed.), *Developmental psycholinguistics: Theory and applications. Georgetown University Round Table on Languages and Linguistics.* Washington, DC: Georgetown University Press.

Bever, T. G. (1982) (Ed.) *Regression in mental development: Basic Processes and mechanisms.* Hillsdale, NJ: Erlbaum.

Bloom, L. (1973) *One word at a time: The use of single word utterances before syntax.* The Hague: Mouton.

Bohannon, J. N. & Marquis, A. L. (1977) Children's control of adult speech. *Child Development, 48*, 1002-1008.

Bowerman, M. (1982) Starting to talk worse: clues to language acquisition from children's late errors. In S. Strauss, ed., *U-shaped behavioral growth.* New York: Academic Press.

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

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

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

Brown, R. (n.d.) Grammars for Adam, Eve and Sarah at Stages I-V. Unpublished manuscripts, Department of Psychology, Harvard University.

Brown, R. & Bellugi, U. (1964) Three processes in the child's acquisition of syntax. In E. Lenneberg (Ed.), *New directions in the study of language.* Cambridge, MA: MIT Press.

Brown, R., and 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.

Brown, R., and McNeill, D. (1966) The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*, 325-337.

Brown, R., Cazden, C., and de Villiers, J. (1971) Directions for scoring fourteen English morphemes obligatory in some contexts. Unpublished manuscript, Department of Psychology, Harvard University.

Bryant, B. & Anisfeld, M. (1969) Feedback versus no-feedback in testing children's knowledge of English pluralization rules. *Journal of Experimental Child Psychology, 8*, 250-255.

Burnham, K. P. and Overton, W. S. (1978) Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika, 65*, 625-33.

Burnham, K. P. and Overton, W. S. (1979) Robust estimation of population size when capture probabilities vary among animals. *Ecology 60*, 927-936.

Bybee, J. L. & Slobin, D. I. (1982a) Rules and schemes in the development and use of the English past tense. *Language, 58*, 265-289.

Bybee, J. L. & Slobin, D. I. (1982b) Why small children cannot change language on their own: Suggestions from the English past tense. In A. Ahlqvist (Ed.), *Papers from the 5th International Conference on Historical Linguistics. (Current Issues in Linguistic Theory Vol. 21, Amsterdam Studies in the theory and history of linguistic science IV.)* Philadelphia/Amsterdam: Benjamins.

Bybee, J. L. (1985) *Morphology: a study of the relation between meaning and form.* Philadelphia: Benjamins.

Carlton, L. E. (1947) Anomalous preterite and past participle forms in the oral language of average fourth grade children. *American Speech, 22*, 40-45.

Carroll, J. B. (1961) Language development in children. In S. Saporta (Ed.), *Psycholinguistics: A book of readings.* New York: Holt Rinehart & Winston.

Carstairs, A. (1987) *Allomorphy in inflexion.* London: Croon Helm.

Cazden, C. B. (1966) Verb inflections. Unpublished manuscript, Graduate School of Education, Harvard University.

Cazden, C. B. (1968) The acquisition of noun and verb inflections. *Child Development, 39*, 433-448.

Cazden, C. B. (1972) *Child language and education.* New York: Holt Rinehart & Winston.

Chamberlain, A. F. (1906) Preterite forms, etc., in the language of English-speaking children. *Modern Language Notes, 21*, 42-44.

Chao, A. (1987) Estimating the population size for capture-recapture data with unequal capturability. *Biometrics, 43*, 783-791.

Chomsky, N. (1959) A Review of B. F. Skinner's "Verbal Behavior". *Language*, *3*, 26-58.

Chomsky, N. & Halle, M. (1968) *The Sound Pattern of English.* New York: Harper and Row.

Clark, E. V. (1982) The young word maker: a case study of innovation in the child's lexicon. In E. Wanner and L. R. Gleitman, eds., *Language acquisition: The state of the art.* New York: Cambridge University Press.

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

Clark, E. V. (1988) On the logic of Contrast. *Journal of Child Language, 15*, 317-335.

Clark, E. V. (1990) On the pragmatics of Contrast. *Journal of Child Language, 17*, 417-431.

Coltheart, M., Patterson, K.E., and Marshall, J.C., ed. 1980. *Deep Dyslexia.* London: Routledge and Kegan Paul.

Cox, M. V. (1989) Children's over-regularization of nouns and verbs. *Journal of Child Language, 16*, 203-206.

de Villiers, J., G., & de Villiers, P. A. (1973) A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research, 2*, 267-278.

Derwing, B. & Baker, W. (1979) Recent research on the acquisition of English morphology. In P. Fletcher & M. Garman (Eds.), *Language Acquisition*. New York: Cambridge University Press.

Edwards, M. L. (1970) An annotated bibliography on the acquisition of English verbal morphology. *Ohio State University Working Papers in Linguistics, 4*, 149-164.

Egedi, D. M. & Sproat, R. W. (1991) Connectionist networks and natural language morphology. Unpublished manuscript, Linguistics Research Department, AT&T Bell Laboratories, Murray Hill, NJ.

Ervin, S. M, & Miller, W. R. (1963) Language development. In H. W. Stevenson (Ed.), *Child psychology: The Sixty-Second Yearbook of the National Society for the Study of Education, Part I.* Chicago: University of Chicago Press.

Ervin, S. M. (1964) Imitation and structural change in children's language. In E. H. Lenneberg (Ed.), *New directions in the study of language*. Cambridge: MIT Press.

Fowler, C. A., Napps, S. E., & Feldman, L. (1985) Relations among regular and irregular morphologically related words in the lexicon as related by repetition priming. *Memory and Cognition, 13*, 241-255.

Francis, N. & Kucera, H. (1982) *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Gasser, M. and Lee, C.-D. (1991). A short-term memory architecture for the learning of morphophonemic rules. In Lippmann, R., Moody, J., and Touretzky, D. (Eds.) *Advances in Neural Information Processing Systems 3*. San Mateo: Morgan Kaufmann.

Gathercole, V. (1979) Birdies like birdseed the bester than buns: A study of relational comparatives and their acquisition. Unpublished doctoral dissertation, University of Kansas.

Gathercole, V. (1987) The contrastive hypothesis for the acquisition of word meaning: A reconsideration of the theory. *Journal of Child Language, 14*, 493-531.

Gathercole, V. (1989) Contrast: A semantic constraint? *Journal of Child Language, 16*, 685-702.

Gelman, D. (1986) The mouths of babes. *Newsweek*, December 15, 84-86.

Gleason J. B. (1980) The acquisition of social speech and politeness formulae. In H. A. R. Giles (Ed.) *Language: Social psychological perspectives*. Oxford: Pergammon.

Gold, E. M. (1967) Language identification in the limit. *Information and Control, 16*, 447-474.

Gordon, P. (1986). Level-ordering in lexical development. *Cognition, 21*, 73-93.

Gordon, P. (1990) Learnability and feedback. *Developmental Psychology 26*, 217-220.

Grimshaw, J., and Pinker, S. (1989) Positive and negative evidence in language acquisition. (Commentary on D. Lightfoot's "The child's trigger experience: 'Degree-0' learnability.") *Behavioral and Brain Sciences, 12*, 341.

Grimshaw, J. & Rosen, S. T. (1990) Knowledge and obedience: The developmental status of the binding theory. *Linguistic Inquiry, 21*, 187-222.

Grudin, J. & Norman, D. A. (1991) Language evolution and human-computer interaction. *Proceedings of the 1991 Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Guillaume, P. (1927) Le developpement des elements formels dans le langage de l'enfant. *Journal de Psychologie, 24*, 203-229. Translated as "The development of formal elements in the child's speech." In C. A. Ferguson & D. I. Slobin (1983) (Eds), *Studies of child language development*. New York: Holt Rinehart & Winston.

Hall, W. S., Nagy, W. E., & Linn, R. (1984). *Spoken words*. Hillsdale, NJ: Erlbaum.

Halle, M. & Mohanan, K. P. (1985) Segmental phonology of modern English. *Linguistic Inquiry, 16*, 57-116.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56*, 208-216.

Higginson, R. (1985) Fixing-assimilation in language acquisition. Unpublished doctoral dissertation, Department of Anthropology, Washington State University.

Horn, E. (1925) Appropriate materials for instruction in reading. In Report of the National Committee on Reading, Whipple, G.M. (Ed.). Twenty-fourth Yearbook of the National Society for the Study of Education, Part I. Bloomington, IL.: Public School Publishing Company.

Horn, M. D. (1927) The thousand and three words most frequently used by kindergarten children. *Childhood Education, 3*, 118-122.

Ingram., D. (1989) First language acquisition: Method, description, and explanation. New York: Cambridge University Press.

Jacobs, R. A., Jordan, M. I., & Barto, A. G (1991) Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science, 15*, 1991.

Karmiloff-Smith, A. & Inhelder, B. (1974/75) If you want to get ahead, get a theory. *Cognition, 3*, 195-212.

Kaye, J. (1990) *Phonology: A cognitive view.* Hillsdale, NJ: Erlbaum Associates.

Kim, J. J., Marcus, G. F., Hollander, & Pinker, S. (1991) Children's inflection is sensitive to morphological structure. *Papers and Reports in Child Language, 30*, 39-46.

Kim, J. J., Pinker, S., Prince, A. Prasada, S. (1991) Why no mere mortal has ever flown out to center field. *Cognitive Science, 15*, 173-218.

Kiparsky, P. (1982). *Explanation in phonology.* Dordrecht: Foris.

Kucera, H. and N. Francis (1967) *Computational analysis of present-day American English.* Providence: Brown University Press.

Kuczaj, S. A. (1976) *-ing, -s, & -ed*: A study of the acquisition of certain verb inflections. Doctoral dissertation, Department of Psychology, University of Minnesota.

Kuczaj, S. A. (1977a) The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior, 16*, 589-600.

Kuczaj, S. A. (1977b) Old and new forms, old and new meanings: The form-function hypothesis revisited. Paper presented at the biennial meeting of the Society for Research in Child Development, New Orleans, March.

Kuczaj, S. A. (1978) Children's judgments of grammatical and ungrammatical irregular past tense verbs. *Child Development, 49*, 319-326.

Kuczaj, S. A. (1981) More on children's initial failure to relate specific acquisitions. *Journal of Child Language, 8*, 485-487.

Lachter, J. & Bever, T. G. (1987) The relation between linguistic structure and associative theories of language learning -- A constructive critique of some connectionist learning models. *Cognition, 28*, 195-247.

Lenneberg, E. H. (1964) A biological perspective on language. In E. H. Lenneberg (Ed.), *New directions in the study of language*. Cambridge: MIT Press.

Lieber, R. (1980) On the organization of the lexicon. Doctoral dissertation, Department of Linguistics and Philosophy, MIT.

Lorge, I. & Chall, J. (1963) Estimating the size of vocabularies of children and adults: An analysis of methodological issues. *The Journal of Experimental Education, 32*, 147-157.

MacKay, D. (1976) On the retrieval and lexical structure of verbs. *Journal of Verbal Learning and Verbal Behavior, 15*, 169-182.

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

MacWhinney, B. (1978) Processing a first language: the acquisition of morphophonology. *Monographs of the Society for Research in Child Development*, 43.

Macwhinney, B. (1990) *The CHILDES Project: Computational Tools for Analyzing Talk (Version 0.88).* Pittsburgh, PA: Department of Psychology, Carnegie-Mellon University.

MacWhinney, B. & Leinbach, J. (in press) Implementations are not conceptualizations: Revising the verb learning model. *Cognition.*

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

MacWhinney, B. & Snow, C. (1990) The Child Language Data Exchange System: an update. *Journal of Child Language, 17*, 457-472.

Maratsos, M. P. (1983) Some current issues in the study of the acquisition of grammar. In P. Mussen, Ed., *Carmichael's Handbook of Child Psychology*, 4th ed. J. H. Flavell and E. H. Markman, Eds., *Volume III: Cognitive Development.* New York: John Wiley & Sons.

Maratsos, M. (1987) Gift of tongues. *Times Literary Supplement* January 2, 19.

Marcus, G. F. (1991) Negative evidence in language acquisition. Submitted for publication.

Marcus, G. F. (in preparation) Children's overregularization of English plurals: A quantitative analysis.

Marchman, V. (1988) Rules and regularities in the acquisition of the English past tense. *Center for Research on Language Newsletter, 2(4)*, University of California, San Diego.

Marchman, V. and Bates, E. (1991) Vocabulary size and composition as predictors of morphological development. Center for Research in Language Technical Report #9103, University of California, San Diego.

McClelland, J. L., Rumelhart, D. E., & Hinton, G. E. (1986) The appeal of parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations.* Cambridge, MA: Bradford Books/MIT Press.

McNeill, D. (1966) Developmental psycholinguistics. In F. Smith and G. Miller, eds., *The Genesis of Language*. Cambridge, Mass.: MIT Press.

Menyuk, P. (1963) A preliminary evaluation of grammatical capacity in children. *Journal of Verbal Learning and Verbal Behavior,* 2, 429-439.

Miller W. R. & Ervin, S. M. (1964) The development of grammar in child language. In U. Bellugi & R. Brown (Eds.), *The acquisition of language. Monographs of the Society for Research in Child Development, 29*, 9-39,

Miller, G. A. (1977) *Spontaneous apprentices -- children and language.* New York, N.Y.: Seabury Press.

Moe, Hopkins, & Rush (1982). The Vocabulary of First-Grade Children. Charles C. Thomas Publisher. Springfield, Ill.

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

Osherson, D. N., Stob, M., and Weinstein, S. (1985) *Systems that learn.* Cambridge, Mass.: MIT Press.

Otis, D. L., Burnham K. P., White G. C. and Anderson, D. R. (1978) Statistical inference from capture data on closed animal populations. *Wildlife Monographs, 63*, 1-135.

Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 93-125). New York, NY: Cambridge University Press.

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, Mass.: MIT Press.

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

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

Pinker, S. (1991) Rules of language. *Science, 253*, 530-535.

Pinker, S., Lebeaux, D. S., and Frost, L. A. (1987) Productivity and constraints in the acquisition of the passive. *Cognition* 26: 195-267.

Pinker, S., and Prince, A. (1988) On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition. *Cognition* 28: 73-193.

Pinker, S., and Prince, A. (1991) Regular and irregular morphology and the psychological status of rules of grammar. *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society.* Berkeley, CA: Berkeley Linguistics Society.

Plunkett, K. & Marchman, V. (1990) From rote learning to system building. Technical Report 9020, Center for Research in Language, University of California, San Diego.

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.

Prasada, S. & Pinker, S. (1991) Generalizations of regular and irregular morphological patterns. Submitted for publication.

Prasada, S. Pinker, S, and Snyder, W. (1990) Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated. Paper presented to the Annual Meeting of the Psychonomic Society, New Orleans, November 23-25.

Prince, A., and Pinker, S. (1988) Rules and connections in human language. *Trends in Neuroscience, 11*, 195-202.

Pyles, T. & Algeo, J. (1982) *The origins and development of the English language.* (Third edition) New York: Harcourt Brace Jovanovich.

Rosenthal, R. (1984) *Meta-analytic procedures for social research.* Newbury Park, CA: Sage Publications, Inc.

Reich, P. A. (1986) *Language development.* Englewood Cliffs, NJ: Prentice-Hall.

Rumelhart, D. E., Hinton, G. E., and Williams, D. (1986) Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations.* Cambridge, MA: Bradford Books/MIT Press.

Rumelhart, D. E. & McClelland, J. L. (1986) On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & The PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models.* Cambridge, MA: Bradford Books/MIT Press.

Rumelhart, D. E. & McClelland, J. L. (1987) Learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Sachs, J. (1983) Talking about there and then: The emergence of displaced reference in parent-child discourse. In K. E. Nelson (Ed.) *Children's Language, Vol 4.* Hillsdale, NJ: Erlbaum.

Sampson, G. (1987) A turning point in linguistics. *Times Literary Supplement*, June 12, 643.

Scalettar, R., and Zee, A. (1988) Emergence of grandmother memory in feed forward networks: learning with noise and forgetfulness. In D. Waltz & J. A. Feldman (Eds.), *Connectionist Models and their Implications: Readings from Cognitive Science*. Norwood, NJ: Ablex.

Seashore, R. H. & Eckerson, L. D. (1940) The measurement of individual differences in general English vocabularies. *Journal of Educational Psychology, 31*, 14-38.

Seber, G. A. F. (1986). A review of estimating animal abundance. *Biometrics, 42*, 267-292.

Slamecka, N. J. (1969) Testing for associative storage in multitrial free recall. *Journal of Experimental Psychology, 81*, 557-560.

Slobin, D. I. (1971) On the learning of morphological rules: A reply to Palermo and Eberhart. In D. I. Slobin (Ed.), *The ontogenesis of grammar: A theoretical symposium*. New York: Academic Press.

Slobin, D. I. (1973) Cognitive prerequisites for the development of grammar. In C. Ferguson and D. I Slobin, ed., *Studies of Child Language Development*. New York: Holt, Rinehart and Winston.

Slobin, D. I. (1978) A case study of early language awareness. In A. Sinclair, R. J. Jarvella, and W. J. M. Levelt (Eds.), *The child's conception of language*. New York: Springer-Verlag.

Slobin. D. I. (1982) Universal and particular in the acquisition of language. In E. Wanner and L. R. Gleitman, ed., *Language Acquisition: the State of the Art*. Cambridge: Cambridge University Press.

Smith, E. E., Langston, C., and Nisbett, R. E. (in press) The case for rules in reasoning. *Cognitive Science*.

Smith, M. E. (1933) Grammatical errors in the speech of preschool children. *Child Development, 4*, 183-190.

Smith, M. E. (1935) A study of some factors influencing the development of the sentence in preschool children. *Journal of Genetic Psychology, 46*, 182-212.

Smolensky, P. (1988) The proper treatment of connectionism. *Behavioral and Brain Sciences, 11*, 1-74.

Stanners, R. F., Neiser, J. J., Hernon, W., P., & Hall, R. (1979) Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior, 18*, 399-41.

Stemberger, J. P. (1989) The acquisition of morphology: analysis of a symbolic model of language acquisition. Unpublished manuscript, Department of Linguistics, University of Minnesota.

Stemberger, J. P. (1982) Syntactic errors in speech. *Journal of Psycholinguistic Research* 11: 313-345.

Stemberger, J. P. & MacWhinney, B. (1986) Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition, 14*, 17-26.

Strauss, S. (Ed.) (1982) *U-shaped behavioral growth*. New York: Academic Press.

Stromswold, K. J. (1990) Learnability and the acquisition of auxiliaries. Doctoral Dissertation, Department of Brain and Cognitive Sciences, MIT.

Talmy, L. (1985) Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen, ed., *Language typology and syntactic description.* Vol. III: *Grammatical categories and the lexicon.* New York: Cambridge University Press.

Templin, M. (1957) *Certain language skills in children.* Minneapolis: University of Minnesota Press.

Ullman, M. & Pinker, S. (1990) Why do some verbs not have a single past tense form? Paper presented to the Fifteenth Annual Boston University Conference on Language Development, October 19-21.

Ullman, M. & Pinker, S. (1991) The English past tense system: Native speaker judgments of regular and irregular forms and the effects of frequency and similarity. Unpublished manuscript, Department of Brain and Cognitive Sciences, MIT.

Valian, V. (in press) Syntactic Subjects in the Early Speech of American and Italian Children. *Cognition, 40.*

Warren-Leubecker A. (1982) Sex differences in speech to children. Master's thesis, Georgia Institute of Technology.

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

Wurzel, W. U. (1989) *Inflectional morphology and naturalness.* Dordrecht: Kluwer.

Zwicky, A. (1970) A double regularity in the acquisition of English verb morphology. *Papers in Linguistics, 3*, 411-418.

Tables

Table 1:
Children Studied

Child	Age	Source	Total Samples	Sampling Frequency
Abe	2;6-5;0	Kuczaj (1977)	210	weekly
Adam	2;3-5;2	Brown (1973)	55	2-3/month
Allison	1;5-2;10	Bloom (1973)	6	occasional
April	1;10-2;11	Higginson (1985)	6	occasional
Eve	1;6-2;3	Brown (1973)	20	2-3/month
Naomi	1;3-4;9	Sachs (1983)	93	weekly to monthly
Nat	2;8	Bohannon (1977)	21	within 1 month
Nathaniel	2;3-3;9	Snow	30	weekly
Peter	1;3-3;1	Bloom (1973)	20	monthly
Sarah	2;3-5;1	Brown (1973)	139	weekly
15 children	4;6-5;0	Hall et al. (1984)	30	2 days/child
24 children	2;1-5;2	Gleason (1980)	72	3 samples/child
20 children	1;6-6;2	Warren-Leubecker (1982)	20	1 sample/child
14 children	2;9-6;6	Gathercole (1979)	16	1-4 samples/child

Table 2:
Overregularization Rates for Individual Children

Child	Correct	Stem+*ed*	Past+*ed*	Total	Overreg Rate
Abe	1786	465	99	2350	0.240
Adam	2444	44	4	2492	0.019
Allison	31	2	0	33	0.061
April	47	6	1	54	0.130
Eve	283	23	1	307	0.078
Naomi	378	34	2	414	0.087
Nat	52	0	0	52	0
Nathan	243	11	3	257	0.054
Peter	853	17	4	874	0.024
Sarah	1717	61	4	1782	0.036
Hall, et al.:					
ANC	79	2	0	81	0.025
BOM	112	1	0	113	0.009
BRD	128	2	0	130	0.015
CHJ	151	4	0	155	0.026
DED	106	5	0	111	0.045
GAT	159	10	0	169	0.059
JOB	130	0	0	130	0
JUB	132	8	0	140	0.057
KIF	100	0	0	100	0
MAA	105	2	0	107	0.019
MIM	77	0	0	77	0
TOS	84	0	0	84	0
TRH	47	3	0	50	0.060
VOH	64	1	0	65	0.015
ZOR	98	0	0	98	0

Mean of Individual Children					.042
Aggregate Databases:					
Gleason	472	32	1	505	0.065
Gathercole	454	16	1	471	0.036
Warren-Leubecker	317	3	0	320	0.009
Mean of Aggregate Databases					.037

Table 3:
Tests of U-Shaped Development:

Correct Irregulars Preceding the Sample with the First Overregularization

Child	**First Overregularization**		**Consecutive Correct in Preceding Samples**	**Overregularization Rate**	**Probability**
	Age	**Sample**			
Adam	2;11	18	381	0.01926	0.0006
Allison	2;10	6	18	0.06061	0.3245
April	2;1	2	3	0.12963	0.6593
Eve	1;8	5	7	0.07818	0.5656
Naomi	1;11	20	15	0.0870	0.2555
Peter	2;6	14	275	0.024027	0.0012
Sarah	2;10	33	231	0.03648	0.0002

Table 4:
Upper-Bound Estimates of Children's Overregularization Rates

Exclusion	**Estimate of Overregularization Rate**		
	Mean	**Median**	**Range**
Samples with < 100 past tokens	0.05	0.03	0.00-0.24
Selected young children			
(Allison, April, Nat, Peter)	0.04	0.025	0.00-0.24
Hall, et al. children	0.07	0.06	0.00-0.24
All 3 Subsets	0.09	0.07	0.02-0.24
Adam, Eve, Sarah, and Peter			
before first overregularization	0.05	0.03	0.00-0.24
Combined with excluded children	0.09	0.07	0.02-0.24
Got (other than Abe)	0.05	0.04	0.00-0.24
Have, be, do (Adam, Eve, Sarah)	0.05	0.03	0.00-0.24
All Four Verbs	0.05	0.04	0.00-0.24
All Exclusions Combined	0.10	0.09	0.04-0.24

Upper Bounds for Individual Children, Following All Exclusions

Child	**Correct**	**Stem+*ed***	**Past+*ed***	**Total**	**Overreg Rate**
Abe	1786	465	99	2350	0.240
Adam	1276	43	4	1323	0.036
Eve	175	21	1	197	0.112
Naomi	281	33	2	316	0.111
Nathan	214	11	3	228	0.061
Sarah	963	61	4	1028	0.063

Table 5:
Growth of Cumulative Regular and Irregular Vocabulary

	Adam	Eve	Sarah
Average Monthly Rate of Increase Before First Overregularization			
Number of Regular Verbs	11.6	12.0	10.2
Proportion of Verbs Regular	2.4	3.6	2.9
Level at Month Before First Overregularization			
Number of Regular Verbs	117	34	68
Number of Irregular Verbs	68	28	53
Total	185	62	121
Percent Regular	63	55	56
Average Monthly Rate of Increase During Overregularization			
Number of Regular Verbs	6.6	11.3	7.2
Proportion of Verbs Regular	0.5	1.1	0.6
Level At End of Transcripts			
Number of Regular Verbs	283	124	269
Number of Irregular Verbs	99	70	99
Total	382	194	368
Percent Regular	74	64	73

Table 6:
Growth of Regular and Irregular Vocabulary, Jackknife Estimates

	Adam	Eve	Sarah
Average Rate of Increase Before First Overregularization			
Number of Regular Verbs	9.1	*	42.2
Proportion of Verbs Regular	-0.2	*	8.3
Level at Month Before First Overregularization			
Number of Regular Verbs	136	*	124
Number of Irregular Verbs	82	*	69
Total	218	*	193
Percent Regular	62	*	64
Average Rate of Growth During Overregularization			
Number of Regular Verbs	9.7	*	7.8
Proportion Regular	1.1	0.6	0.4
Level At End of Transcripts			
Number of Regular Verbs	213	106	209
Number of Irregular Verbs	87	78	96
Total	300	184	305
Percent Regular	71	58	69

*Eve had too few samples before her first overregularization to yield Jackknife estimates.

Table 7:
Proportion of Verb tokens in obligatory past tense contexts that were marked for tense

before the first month with an overregularization

and in the period beginning with the first month.

	Irregular		Regular	
	Proportion Marked	**Number of Obligatory Contexts**	**Proportion Marked**	**Number of Obligatory Contexts**
Adam, Before	.74*	369	.08*	123
Adam, During	.91*	559	.73*	173
Eve, Before	.18*	11	.11*	19
Eve, During	.62*	259	.66*	118
Sarah, Before	.65*	131	.44	26
Sarah, During	.90*	508	.85*	134

Proportions significantly different from .5 are marked with an asterisk.

Figures

Figure 1. Histogram of overregularization rates across 25 children.

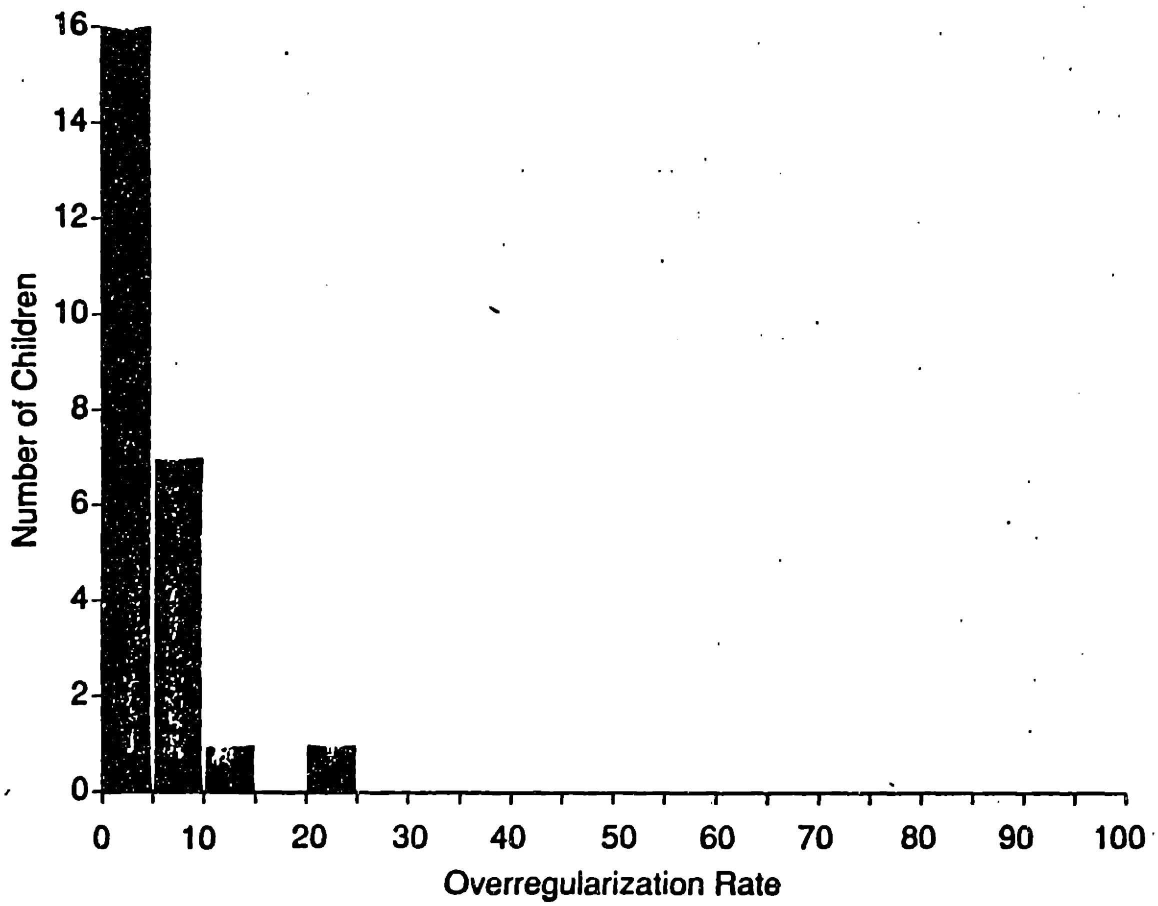


Figure 2. Hypothetical developmental sequence that would yield low overregularization rates as an averaging artifact.

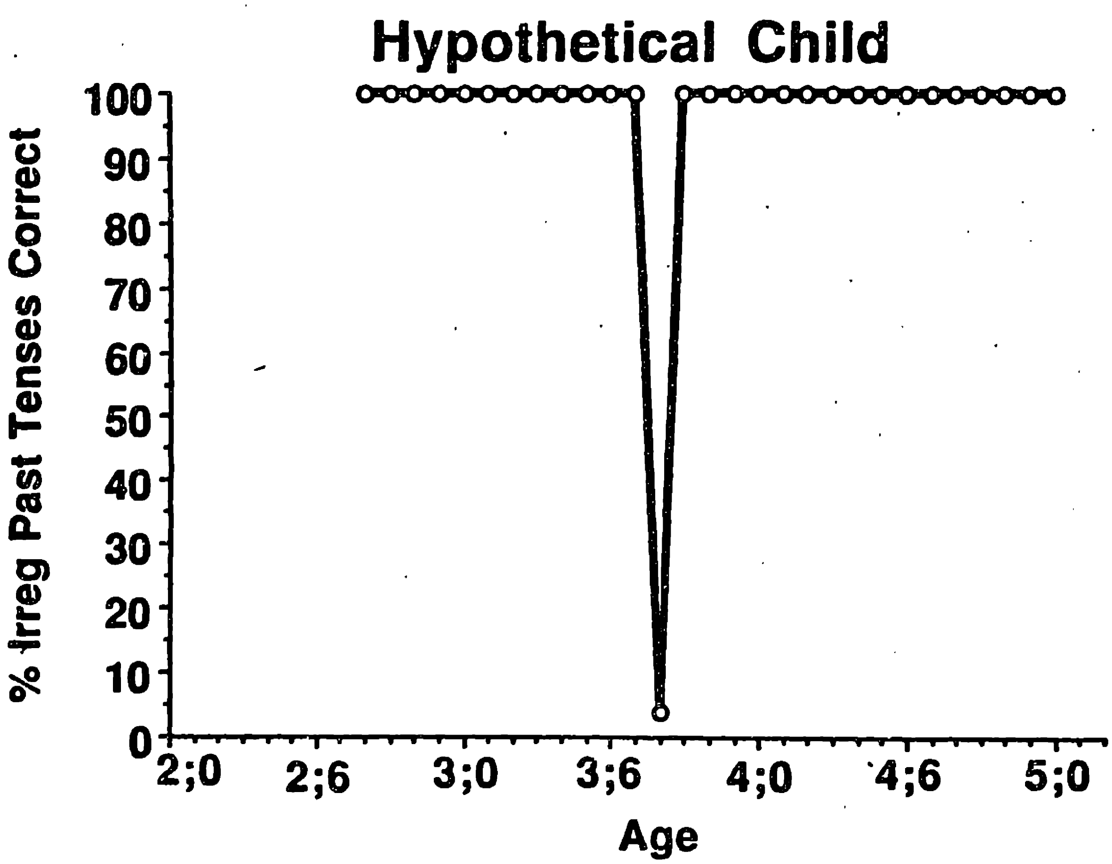


Figure 3. Percentage of Adam's irregular past tense forms that are correct (100% - overregularization rate).

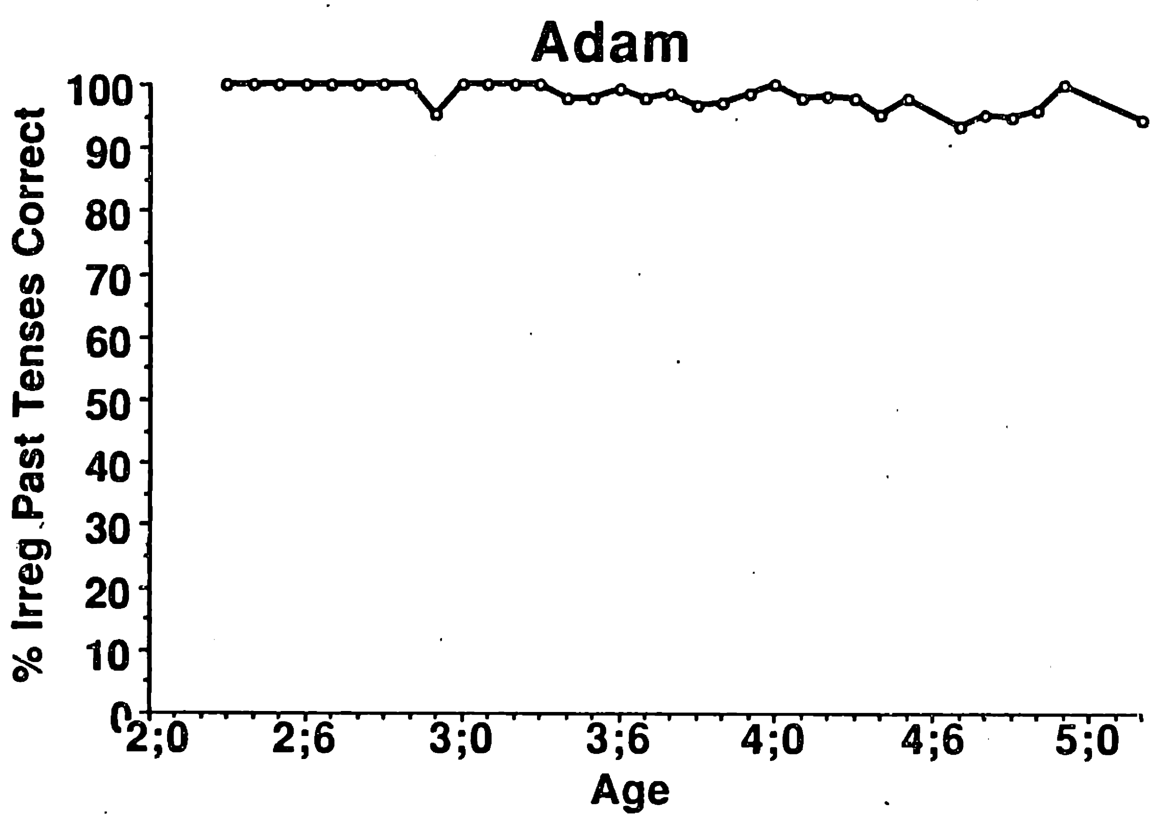


Figure 4. Percentage of Eve's irregular past tense forms that are correct (100% - overregularization rate).

Figure 5. Percentage of Sarah's irregular past tense forms that are correct (100% - overregularization rate).

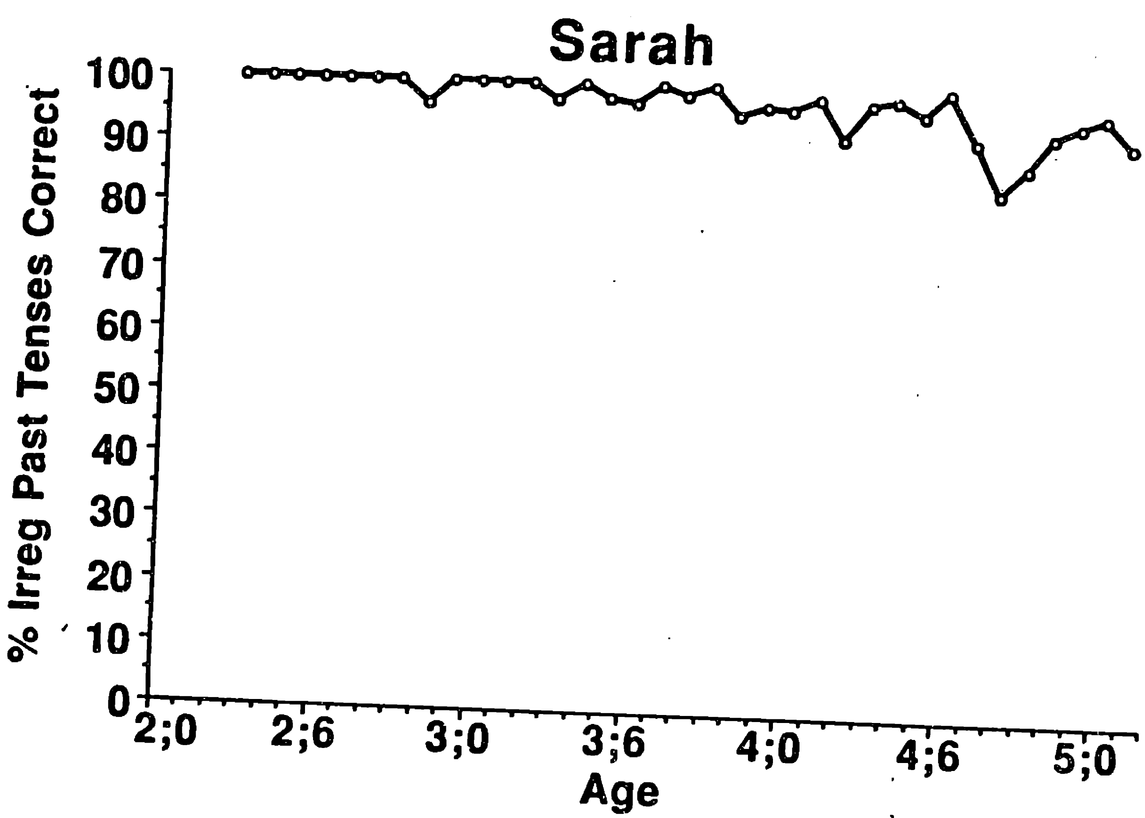


Figure 6. Percentage of Abe's irregular past tense forms that are correct (100% - overregularization rate).

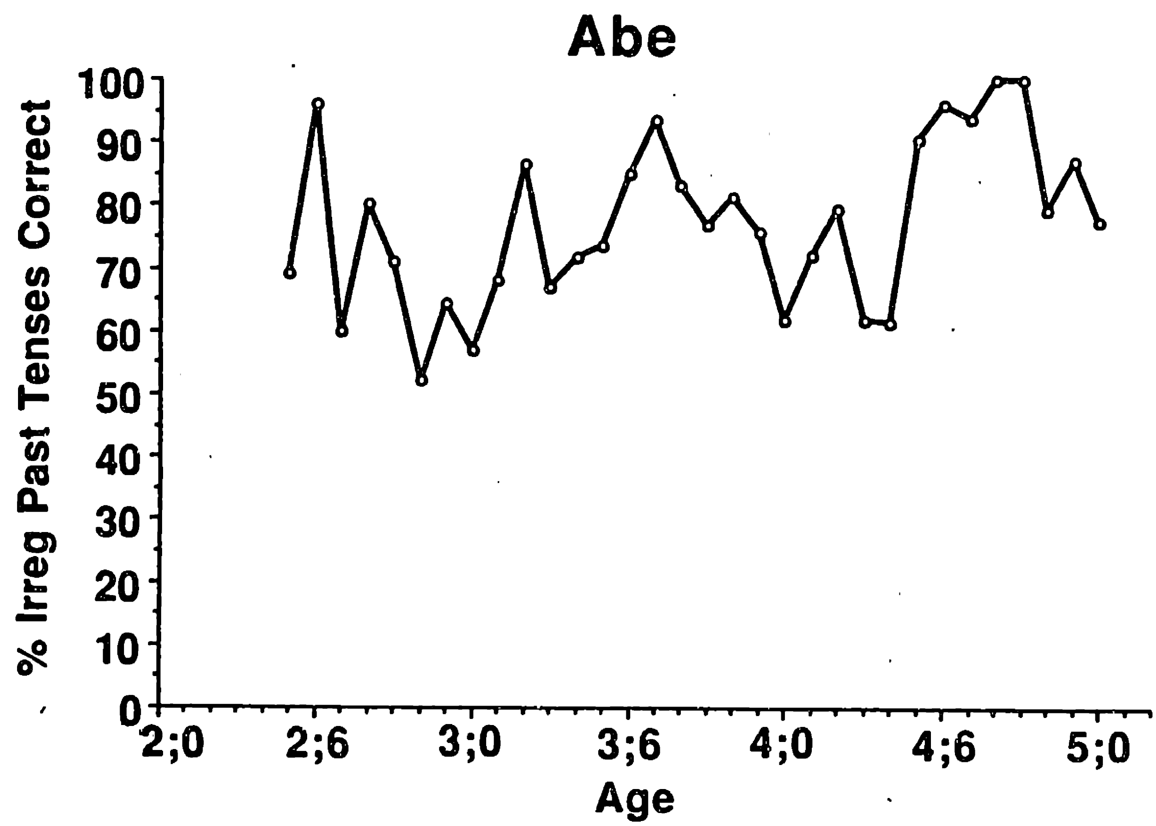


Figure 7. Histogram of overregularization rates of Adam's verbs (10 or more tokens per verb).

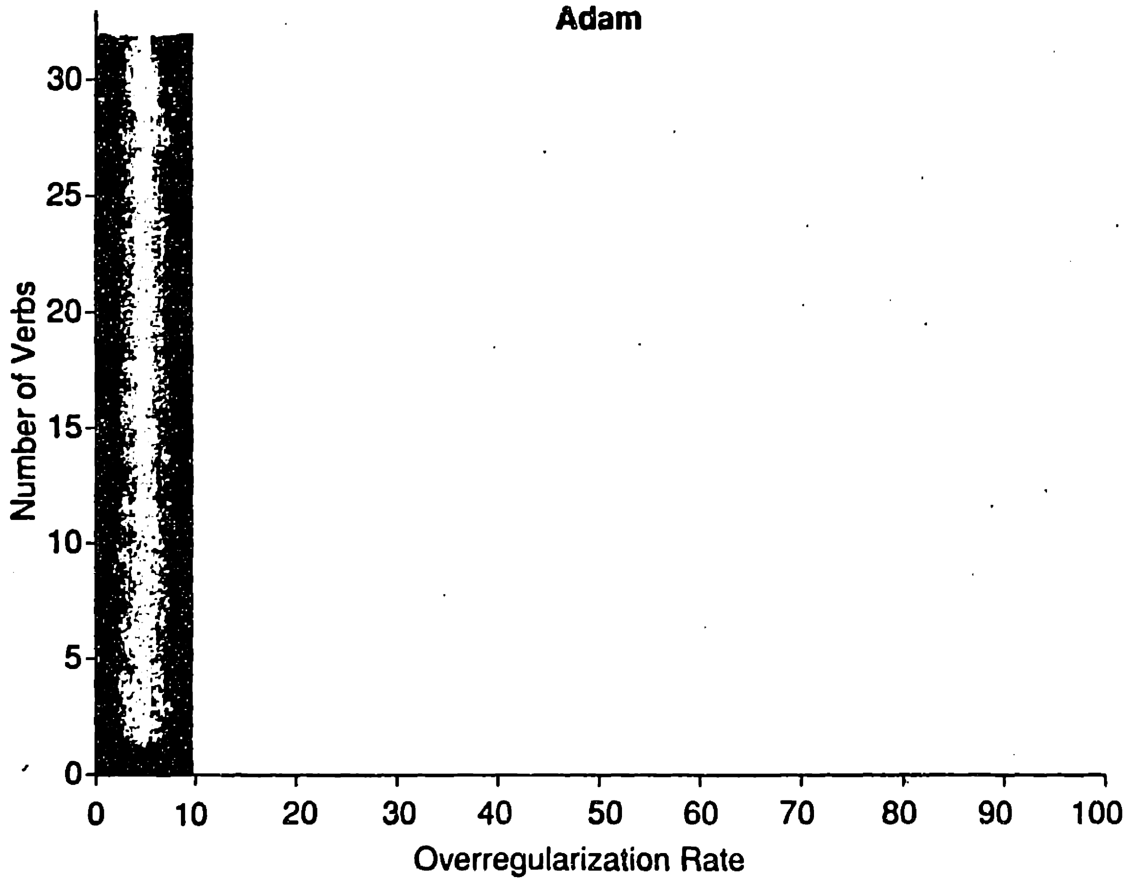


Figure 8. Histogram of overregularization rates of Eve's verbs (10 or more tokens per verb).

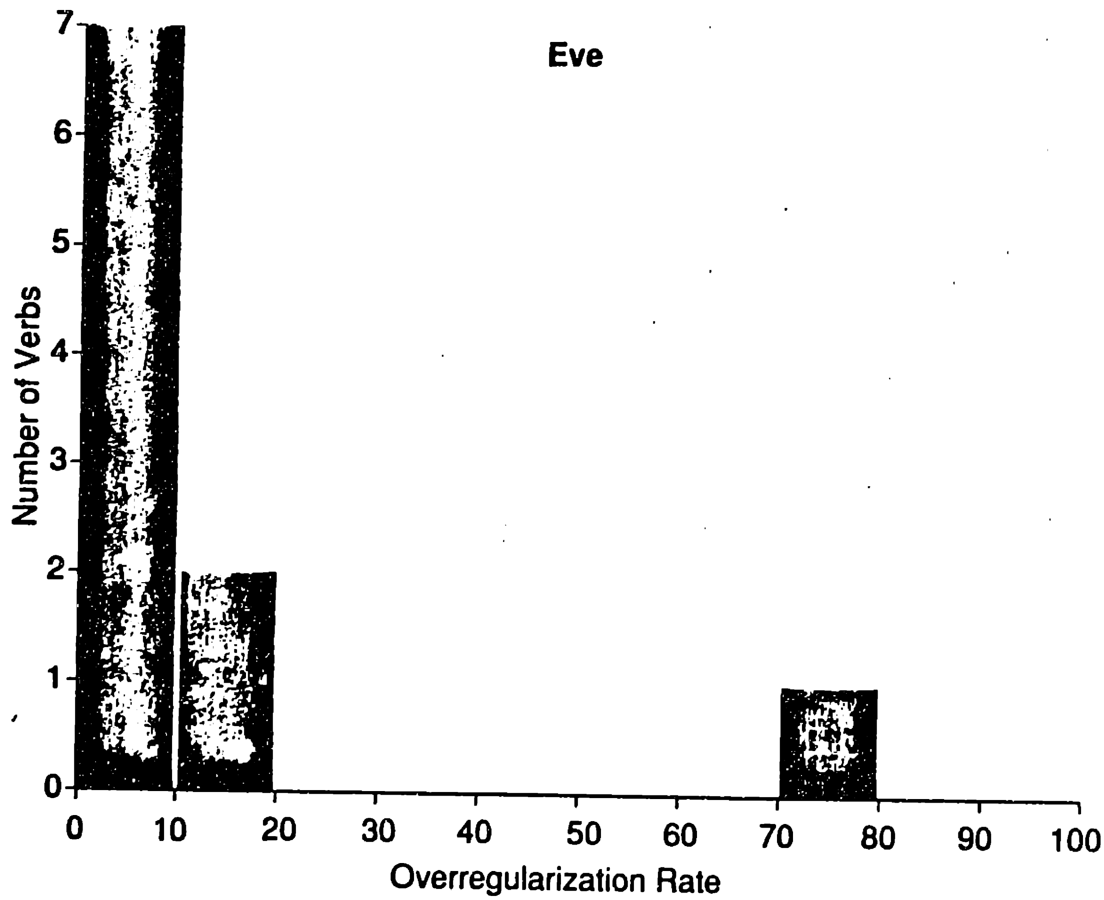


Figure 9. Histogram of overregularization rates of Sarah's verbs (10 or more tokens per verb).

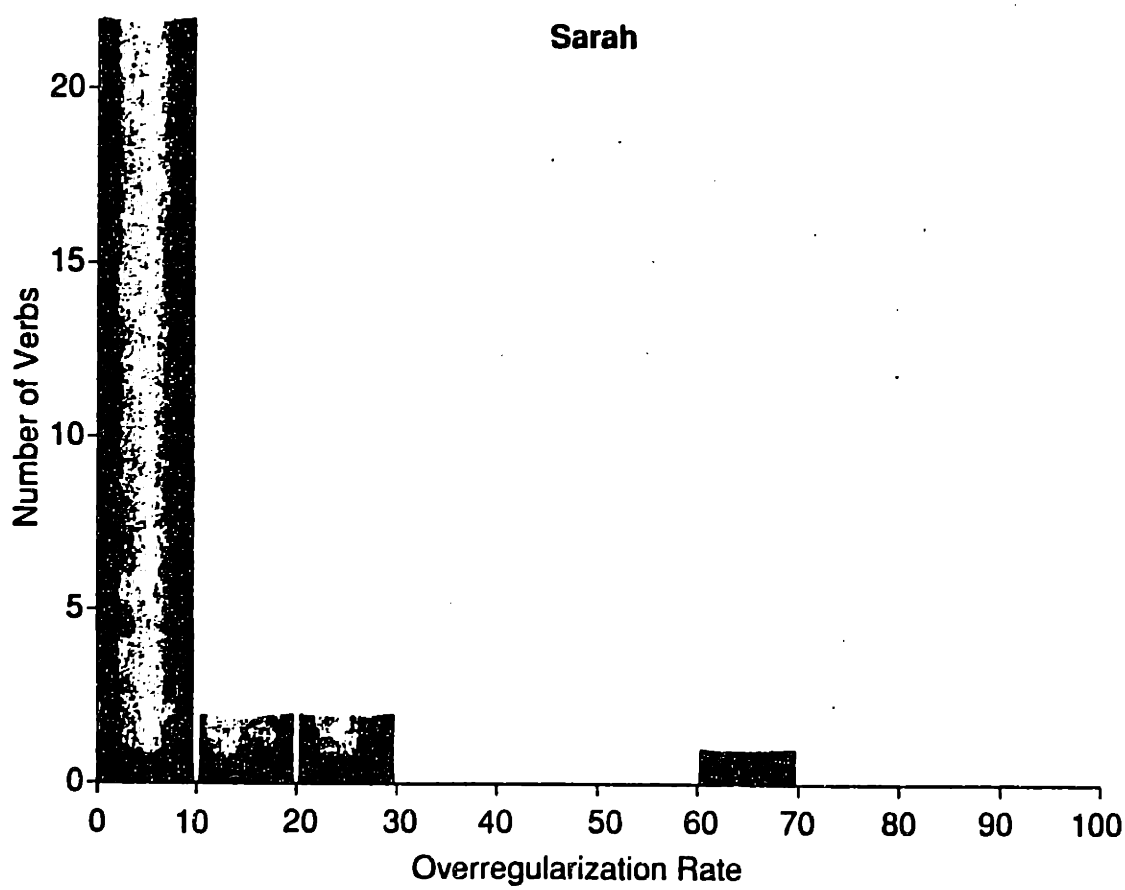


Figure 10. Histogram of overregularization rates of Abe's verbs (10 or more tokens per verb).

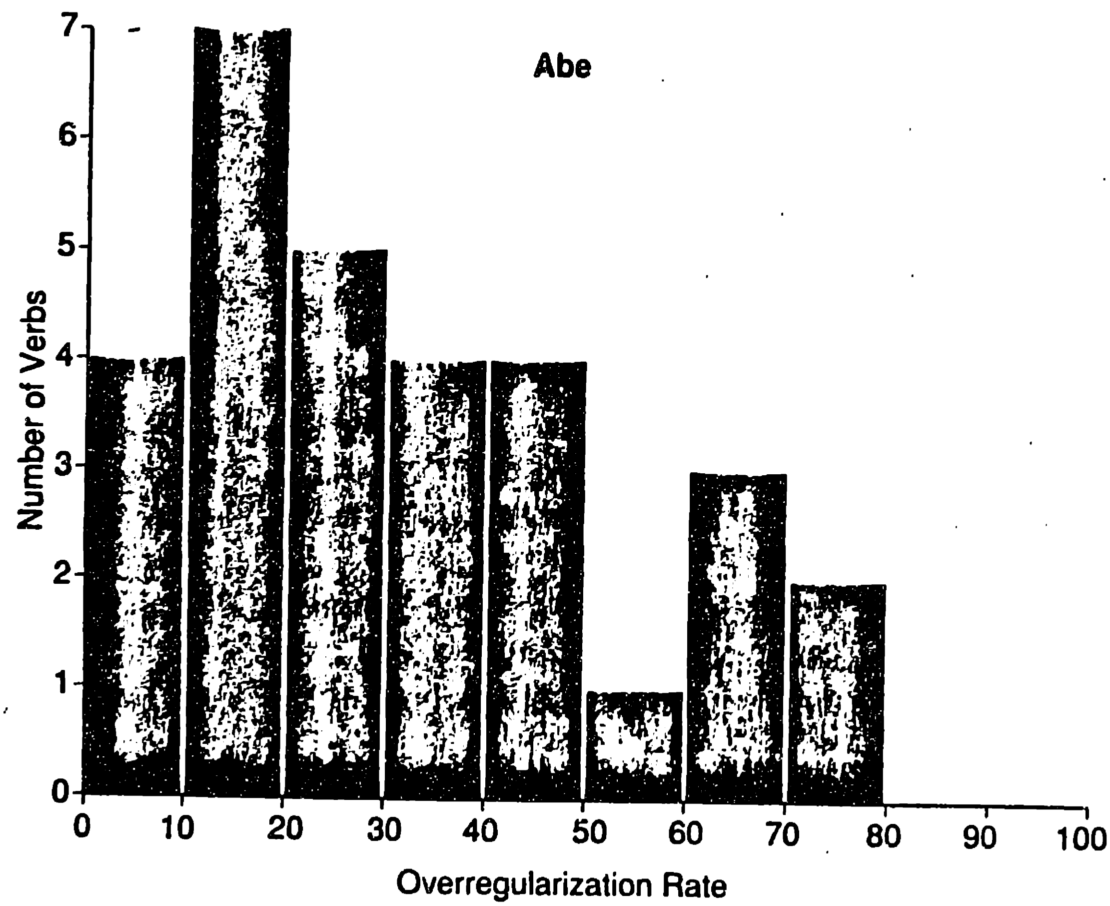


Figure 11. Hypothetical developmental sequence for different verbs that would yield low overregularization rates as an averaging artifact.

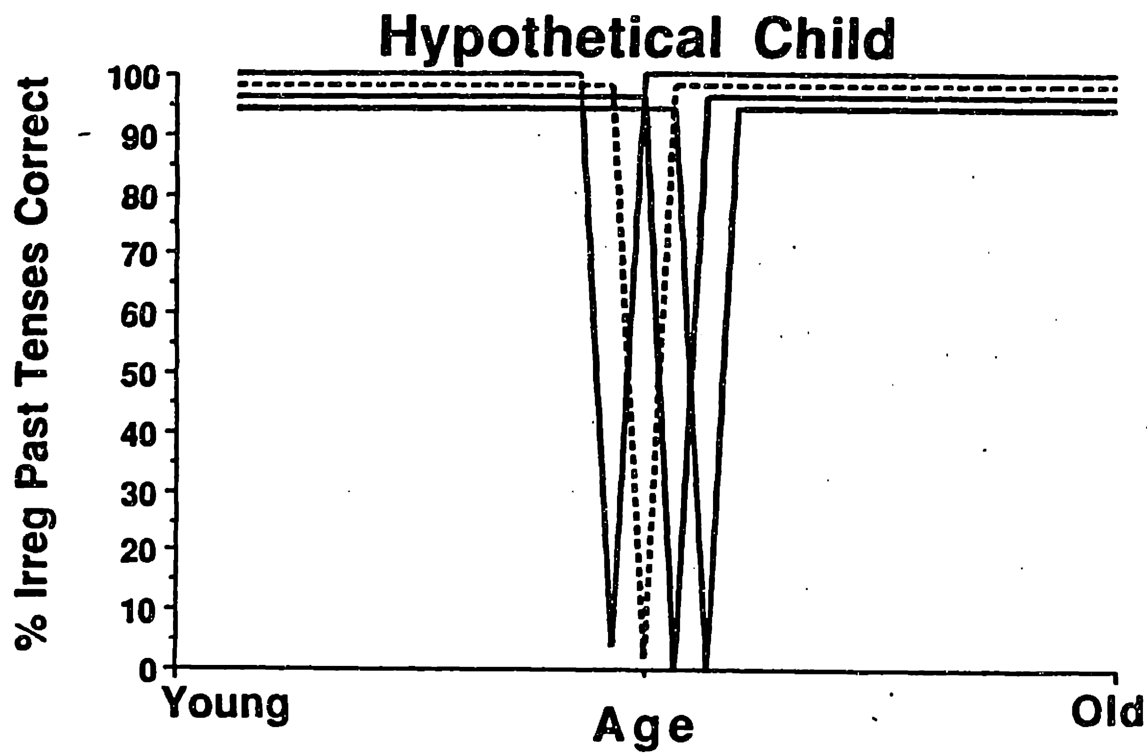


Figure 12. Example of a verb that resists overregularization as the child gets older.

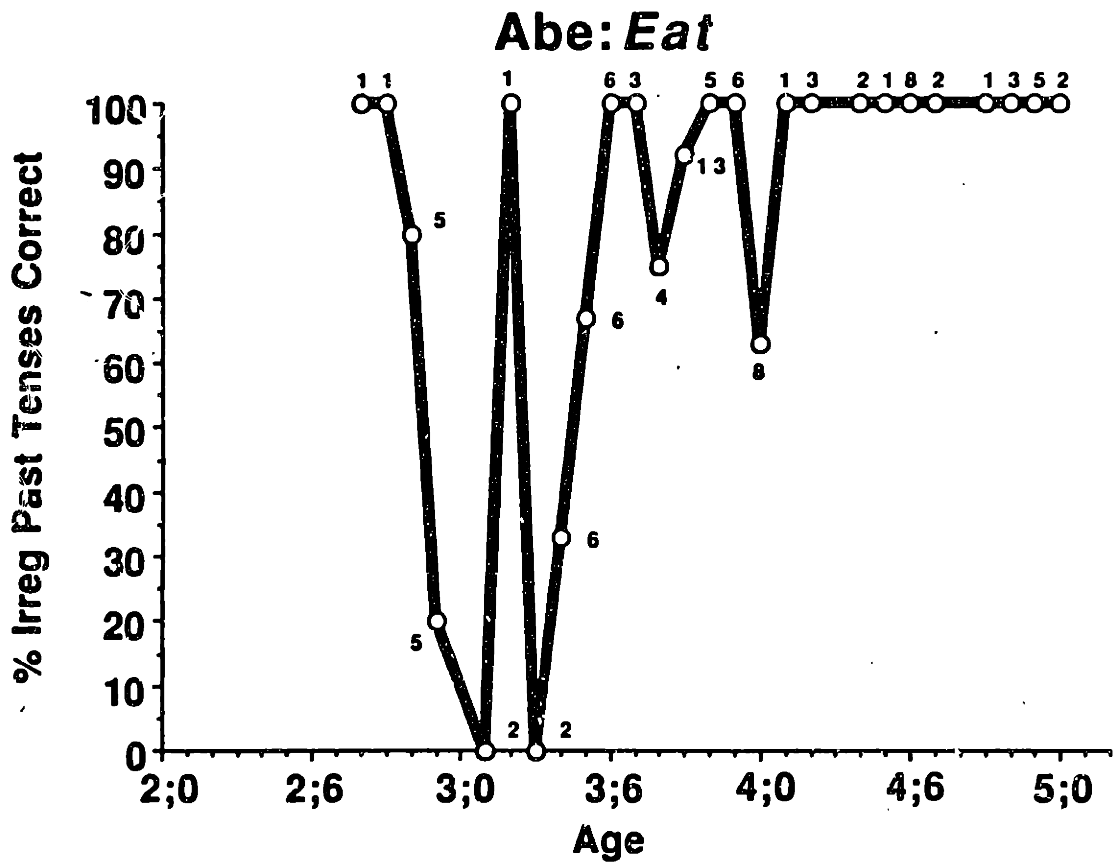


Figure 13. Example of a verb that is rarely overregularized at any age.

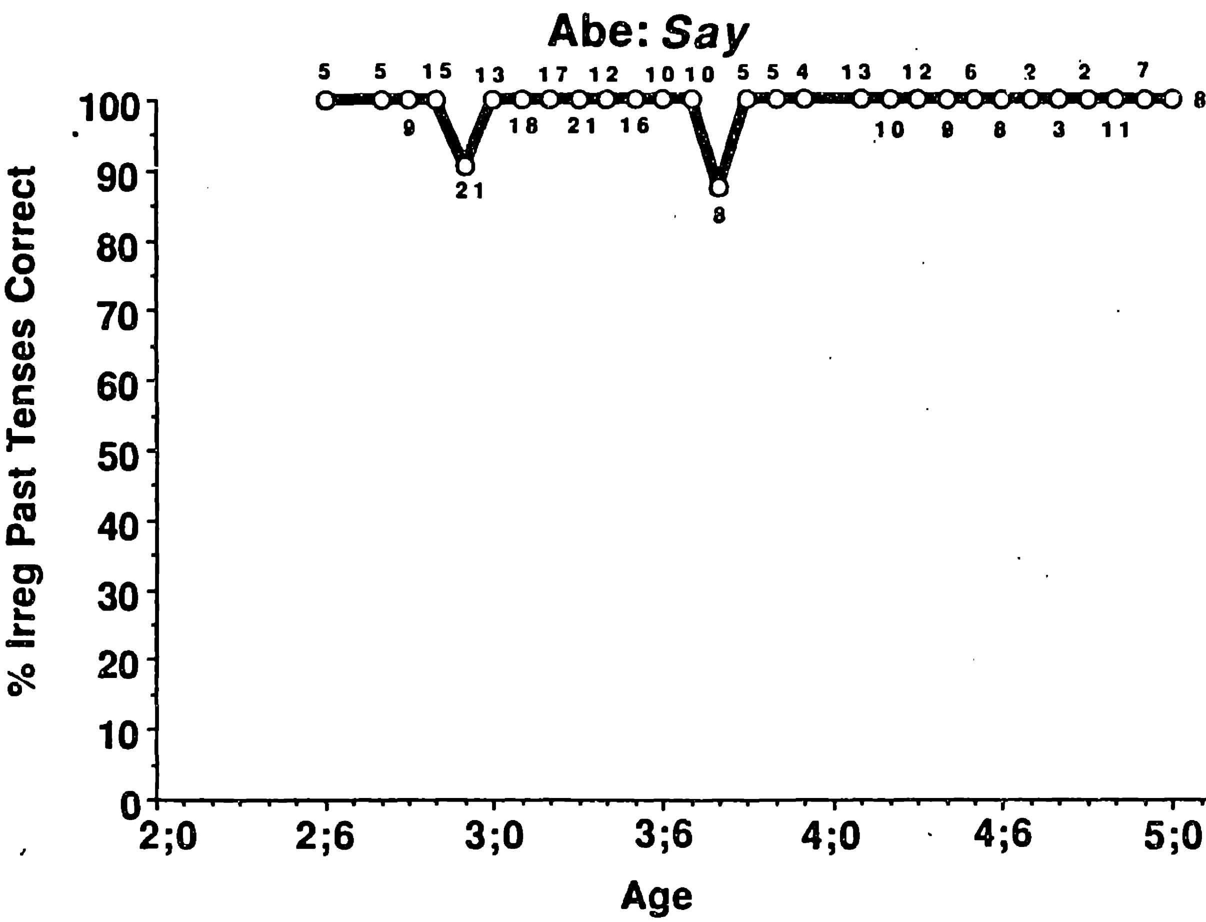


Figure 14. Example of a verb that is overregularized throughout development.

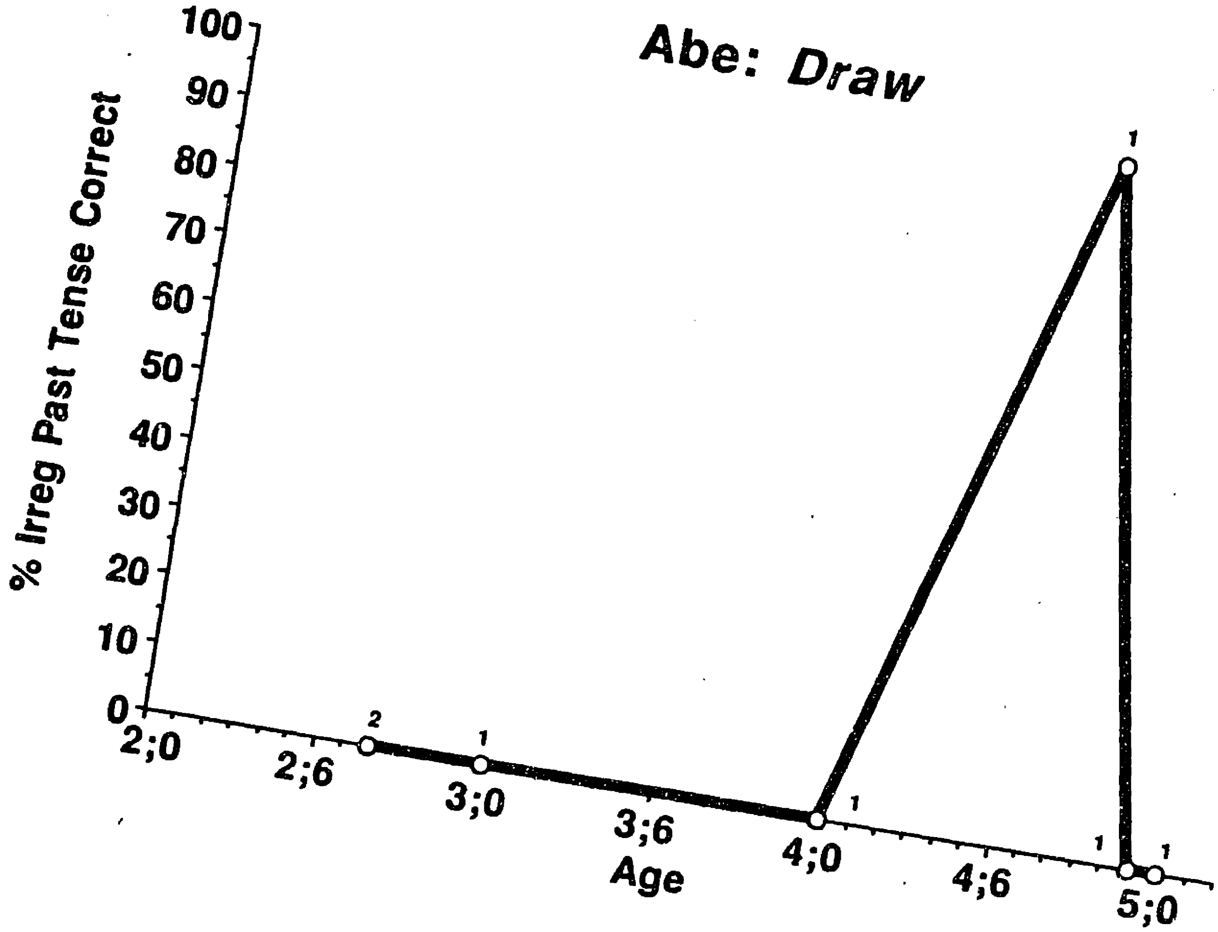


Figure 15. Example of a verb with a chaotic developmental pattern.

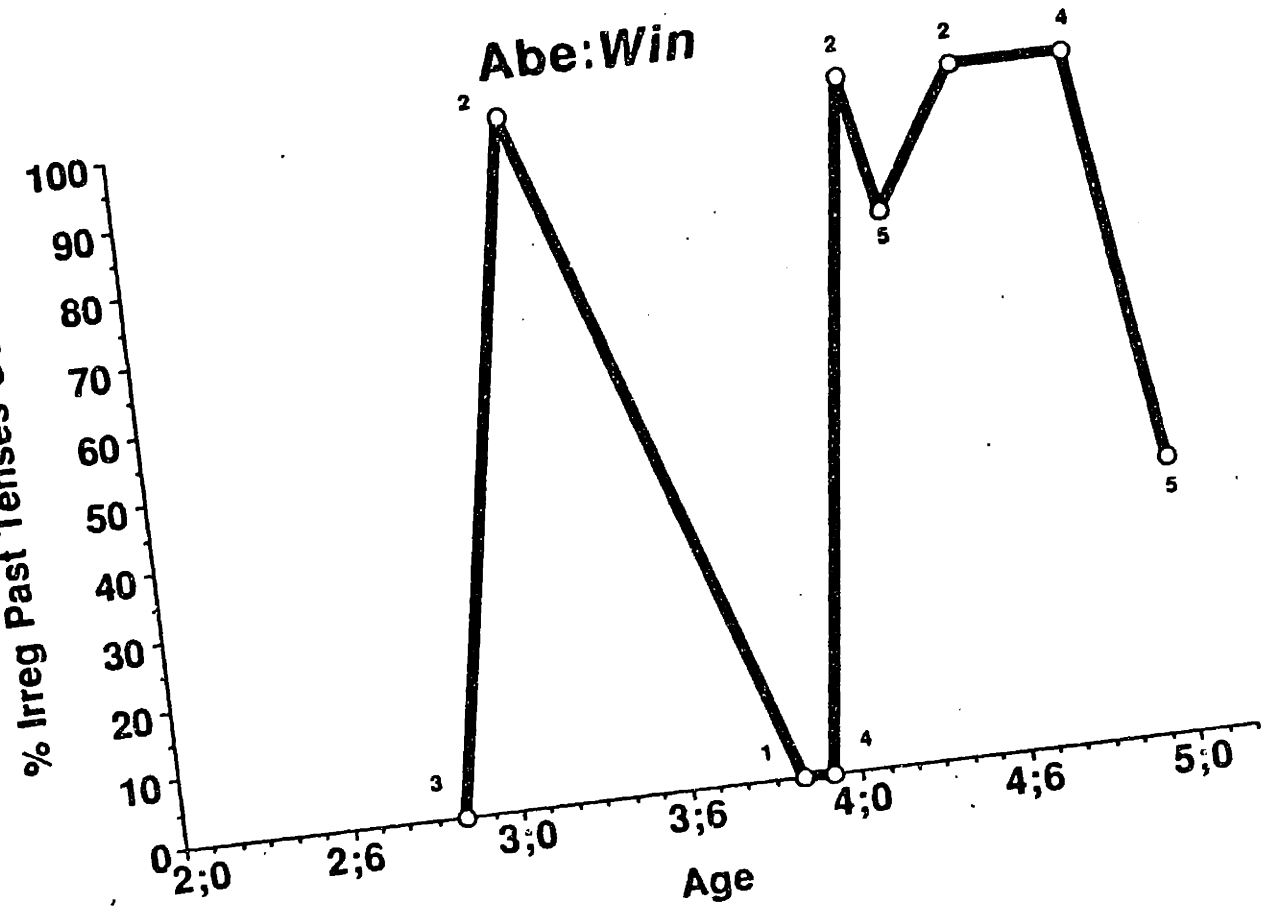


Figure 16. Performance of the Rumelhart-McClelland model on regular and irregular verbs as a function of training epochs. The dip in the curve for irregular verbs following the 10th epoch corresponds to the onset of overregularization. From Rumelhart & McClelland (1986).

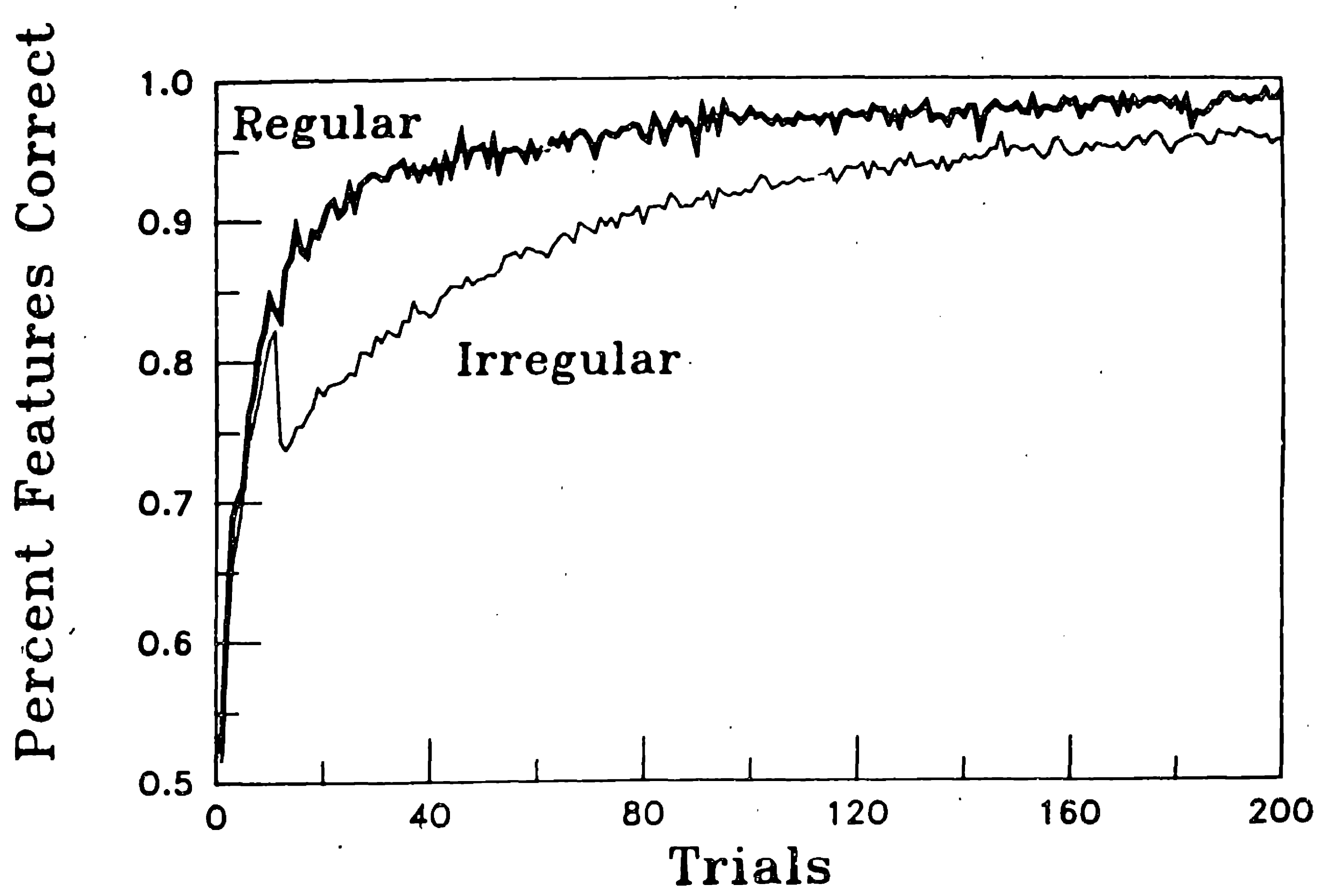


Figure 17. Tendency of the Rumelhart-McClelland model to overregularize irregular verbs as a function of training epochs. Overregularization tendency is measured as the ratio of the strength of the correct irregular response to the sum of the strengths of the correct and the overregularized responses. Points below the line correspond to a tendency to overregularize. From Rumelhart & McClelland (1986).

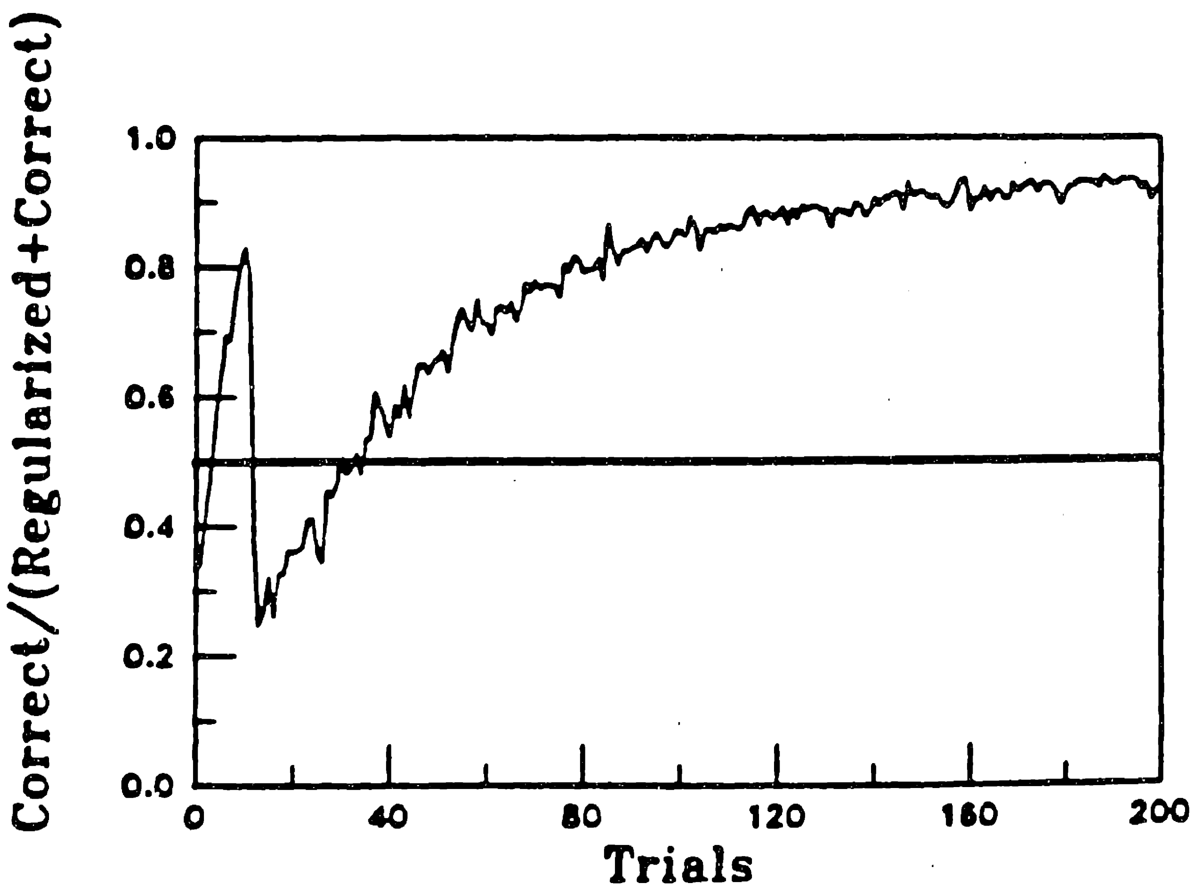


Figure 18. Percentage of verb tokens that are regular for Adam and the adults conversing with Adam.

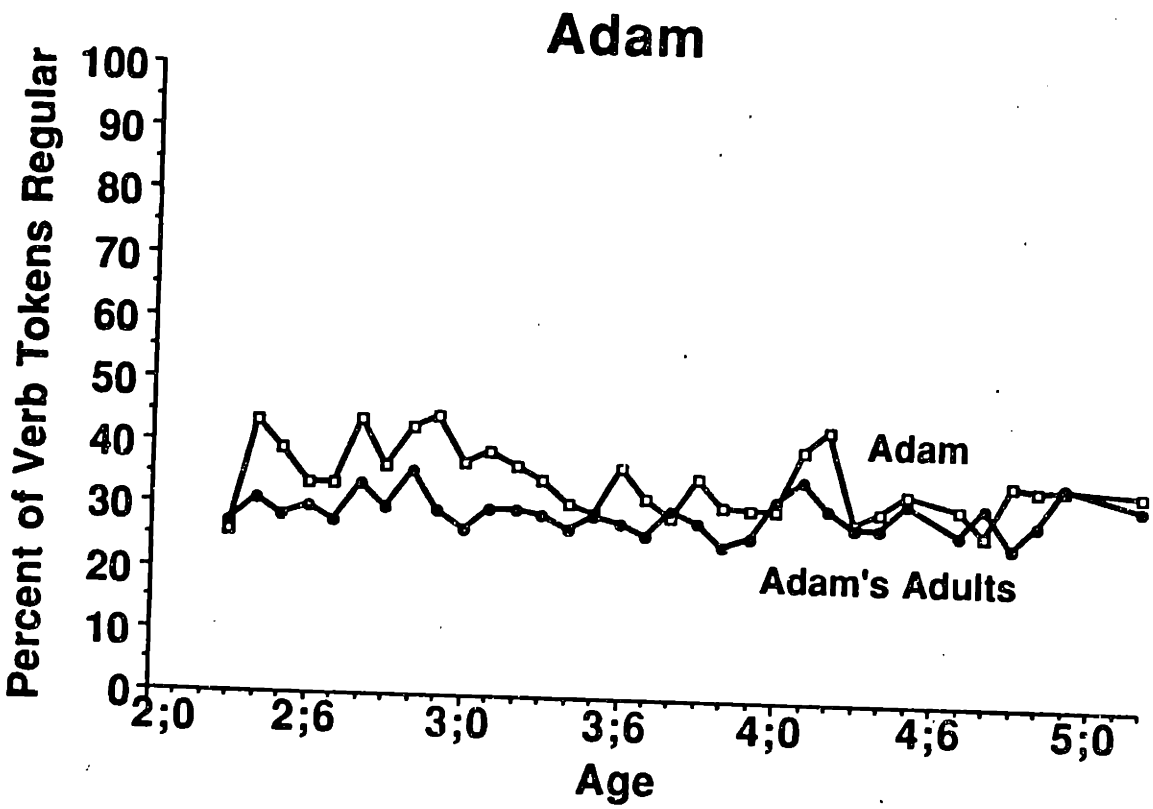


Figure 19. Percentage of verb tokens that are regular for Eve and the adults conversing with Eve.

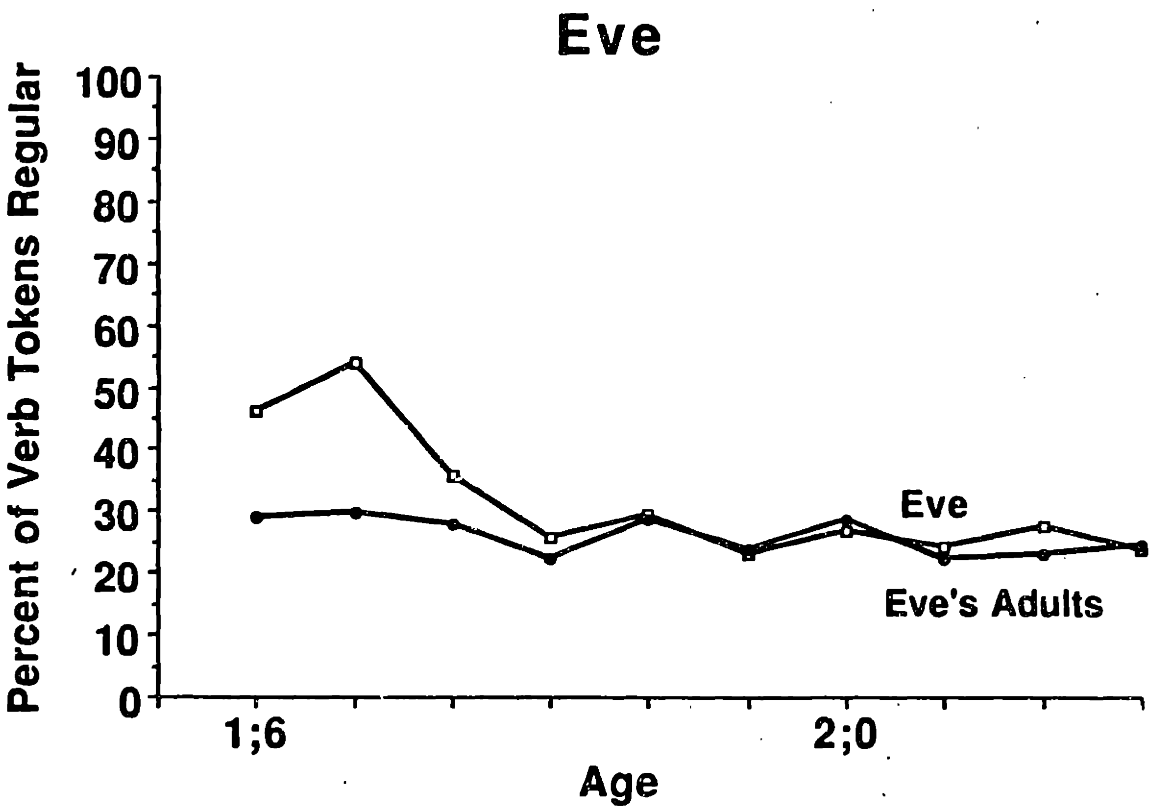


Figure 20. Percentage of verb tokens that are regular for Sarah and the adults conversing with Sarah.

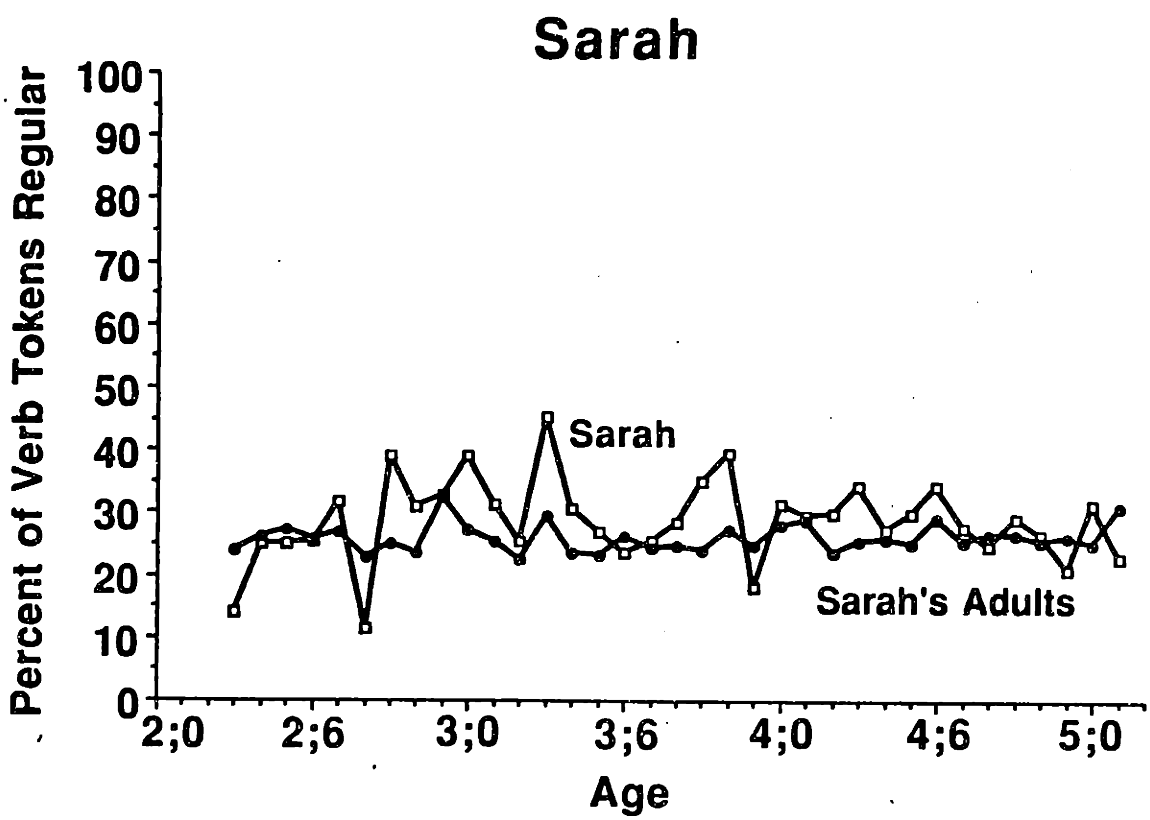


Figure 21. Adam's cumulative regular and irregular verb vocabulary.

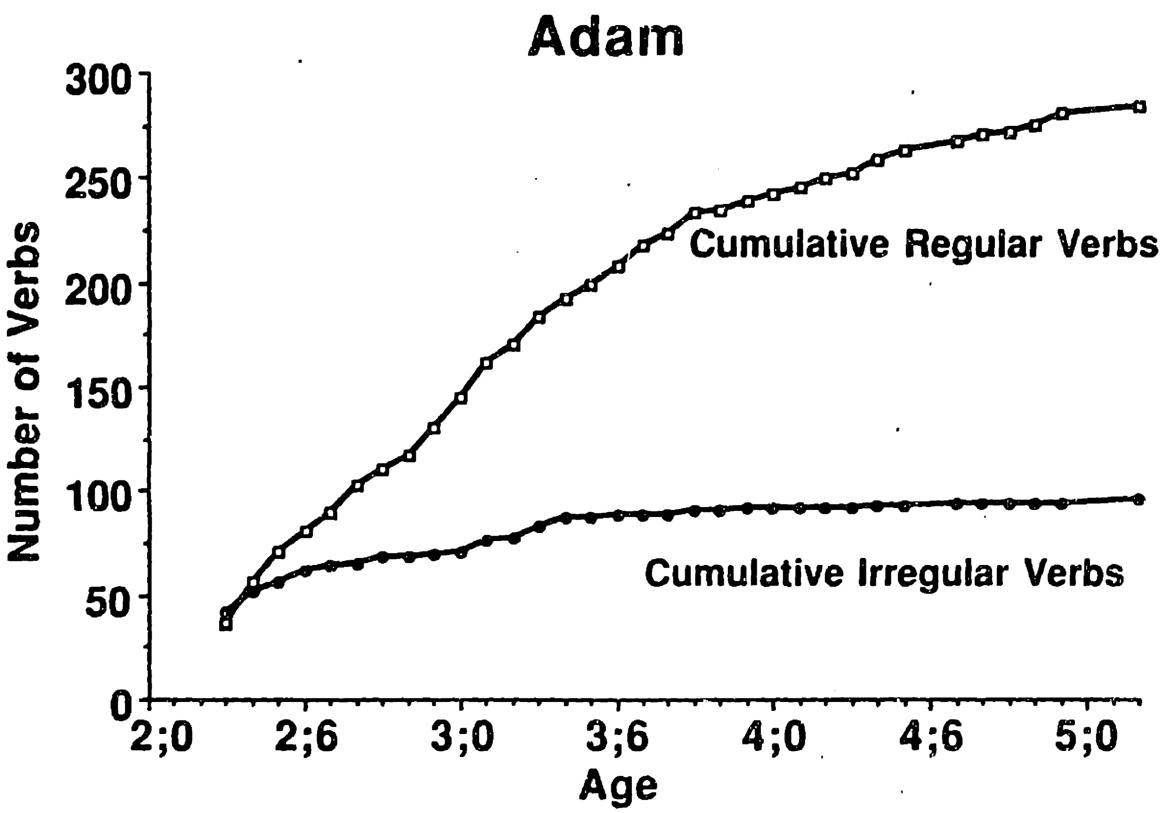


Figure 22. Proportion of Adam's cumulative verb vocabulary that is regular, and his overregularization rate (subtracted from 100%).

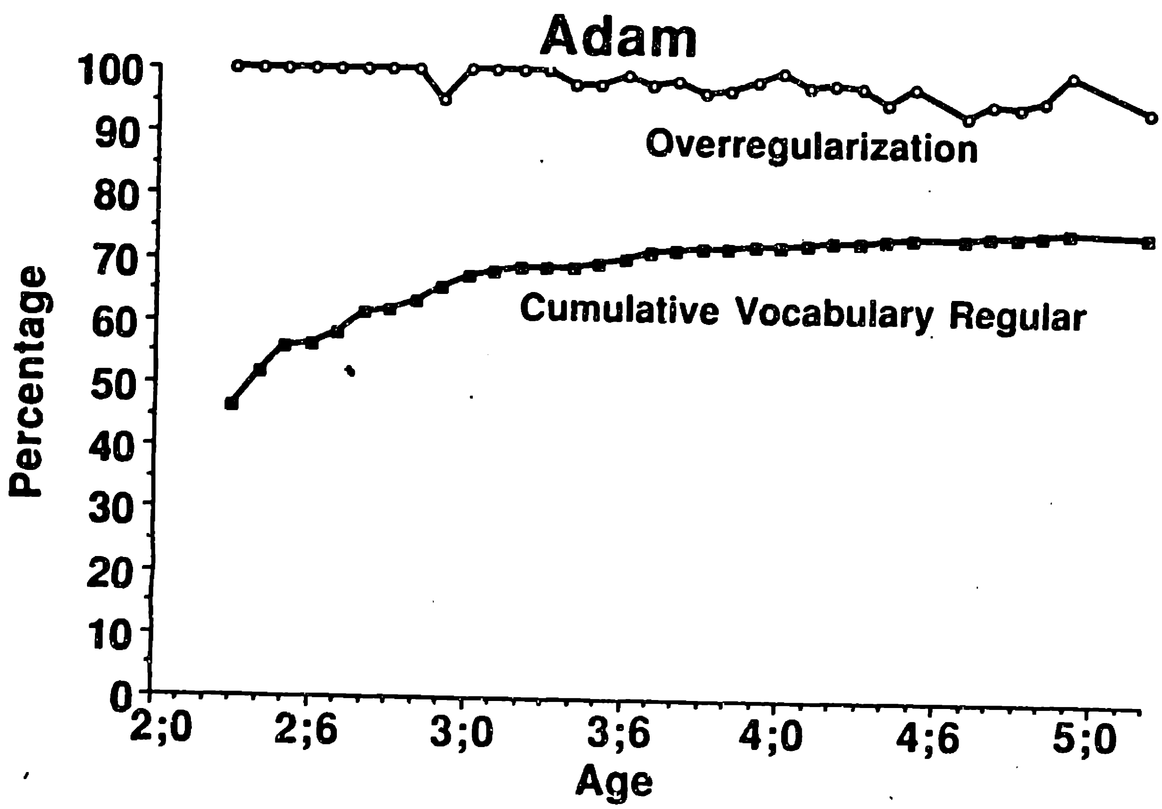


Figure 23. Eve's cumulative regular and irregular verb vocabulary.

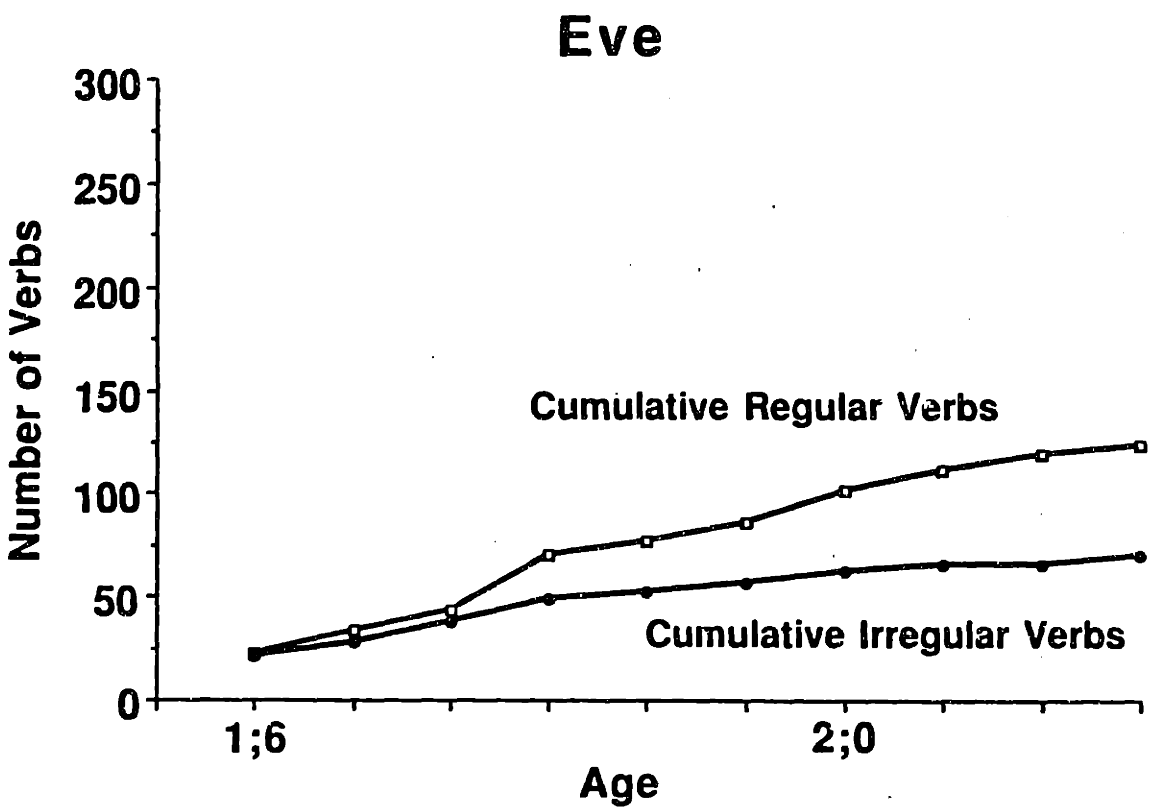


Figure 24. Proportion of Eve's cumulative verb vocabulary that is regular, and her overregularization rate (subtracted from 100%).

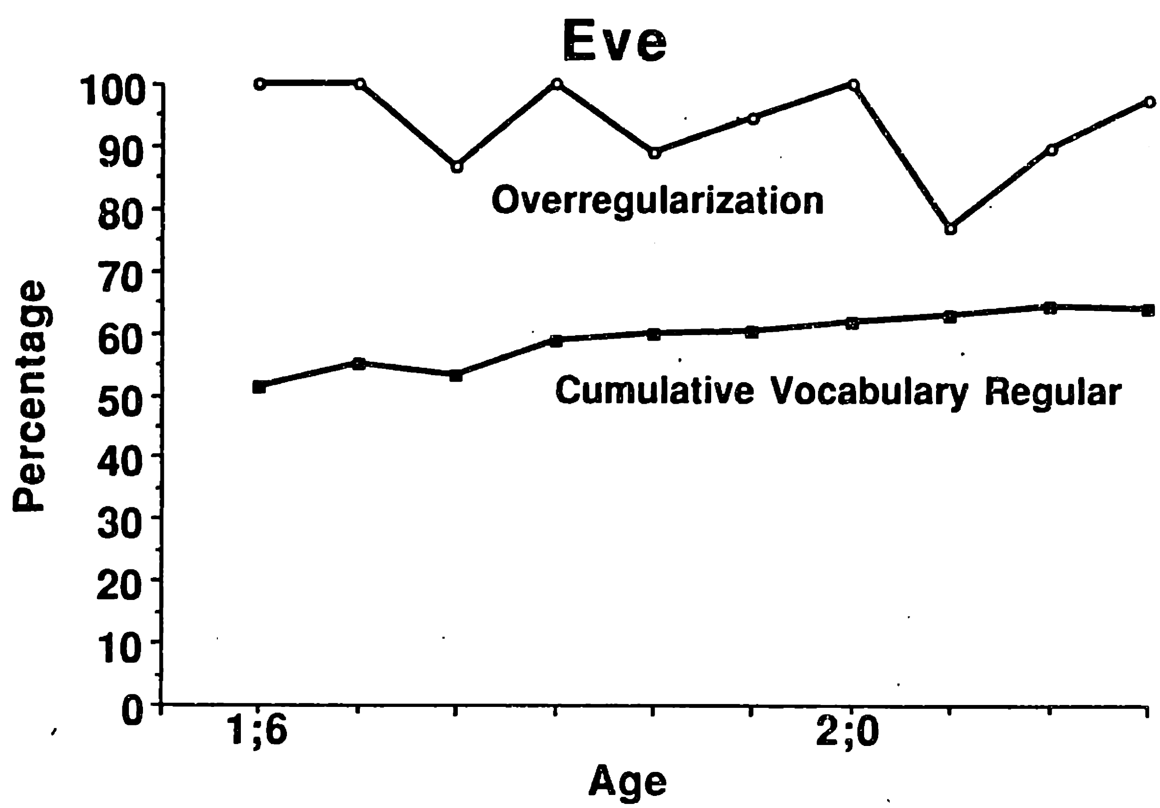


Figure 25. Sarah's cumulative regular and irregular verb vocabulary.

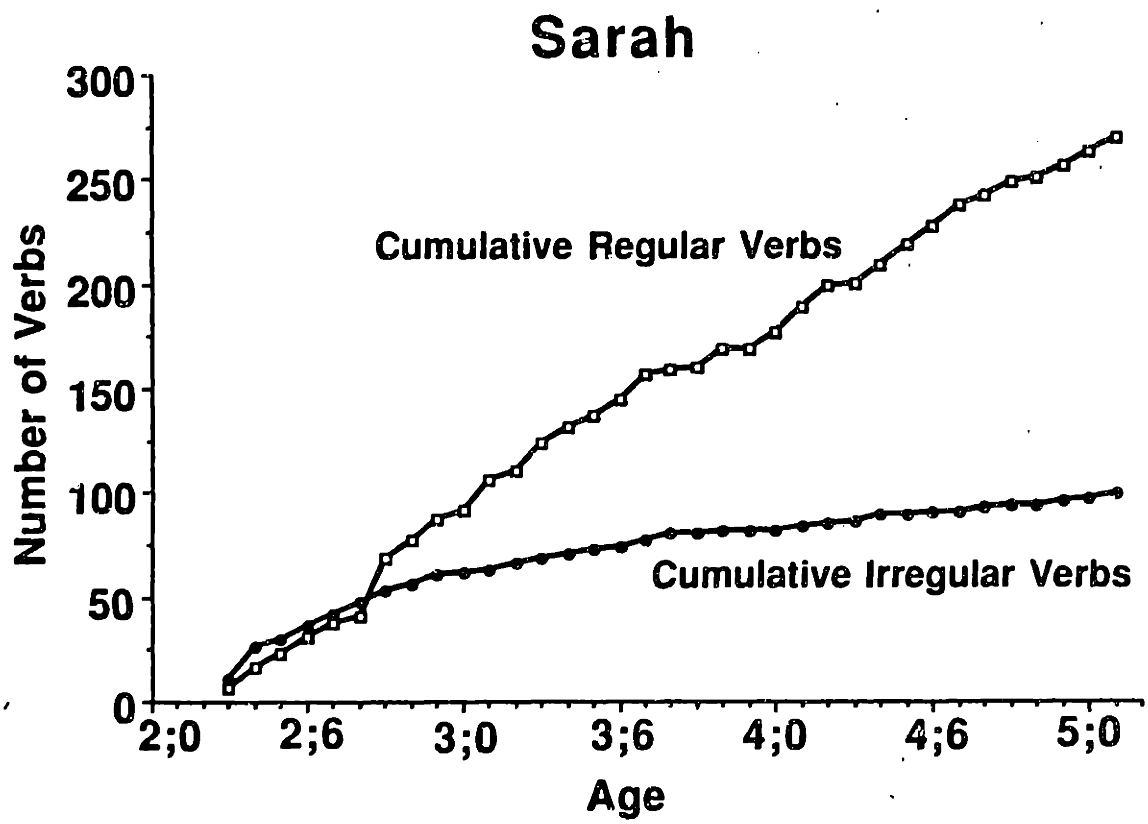


Figure 26. Proportion of Sarah's cumulative verb vocabulary that is regular, and her overregularization rate (subtracted from 100%).

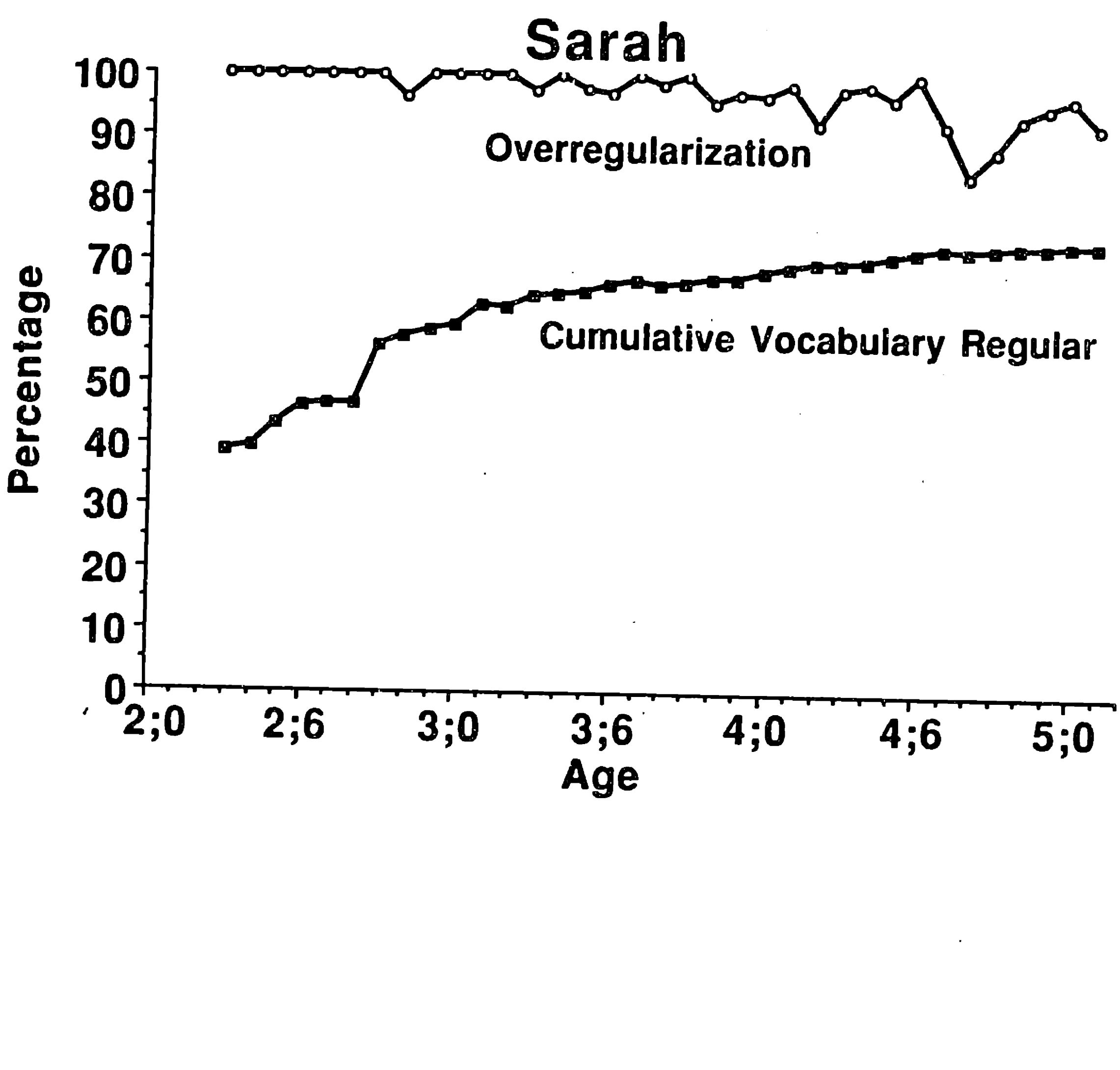


Figure 27. Adam's Jackknife-estimated regular and irregular verb vocabulary.

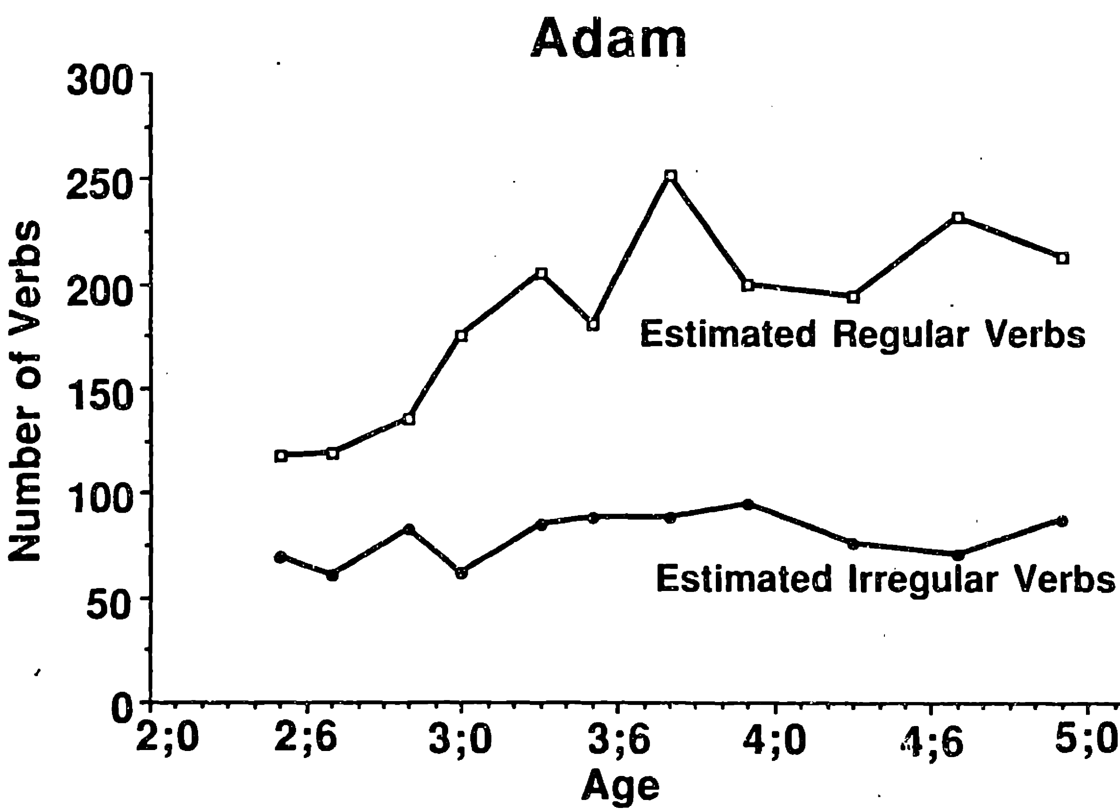


Figure 28. Proportion of Adam's Jackknife-estimated verb vocabulary that is regular, and his overregularization rate (subtracted from 100%).

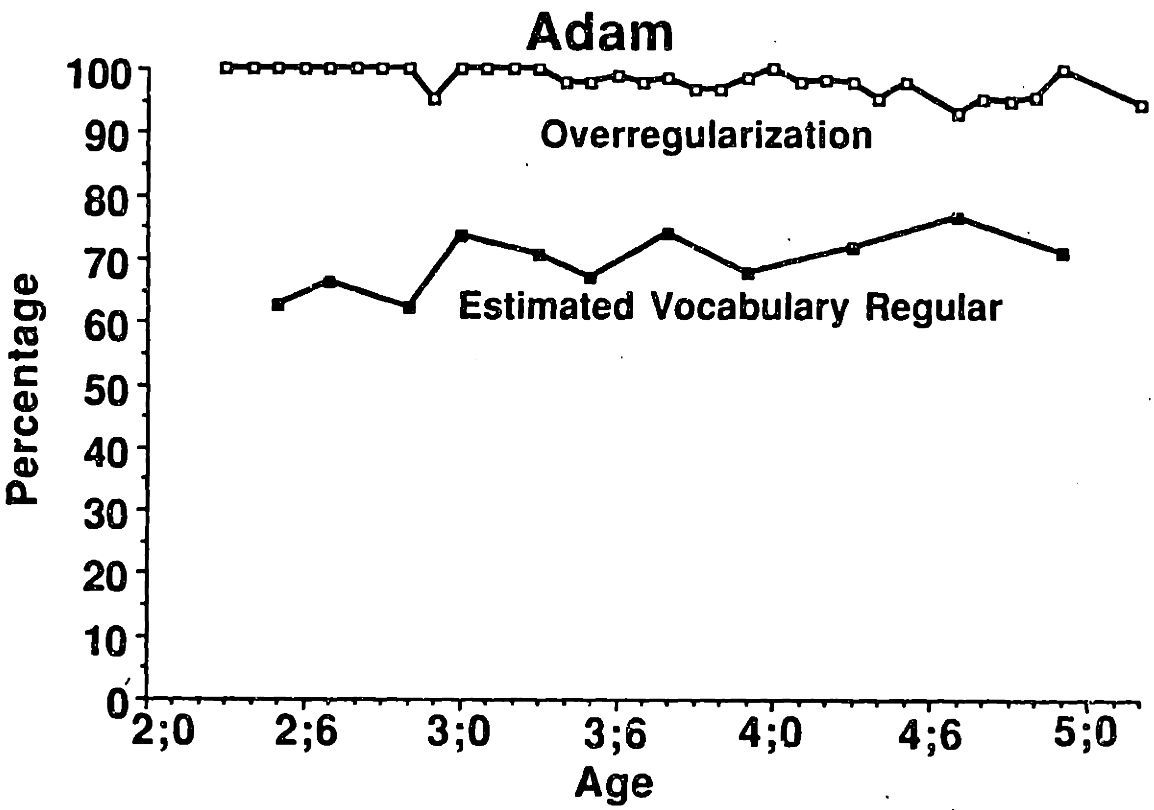


Figure 29. Eve's Jackknife-estimated regular and irregular verb vocabulary.

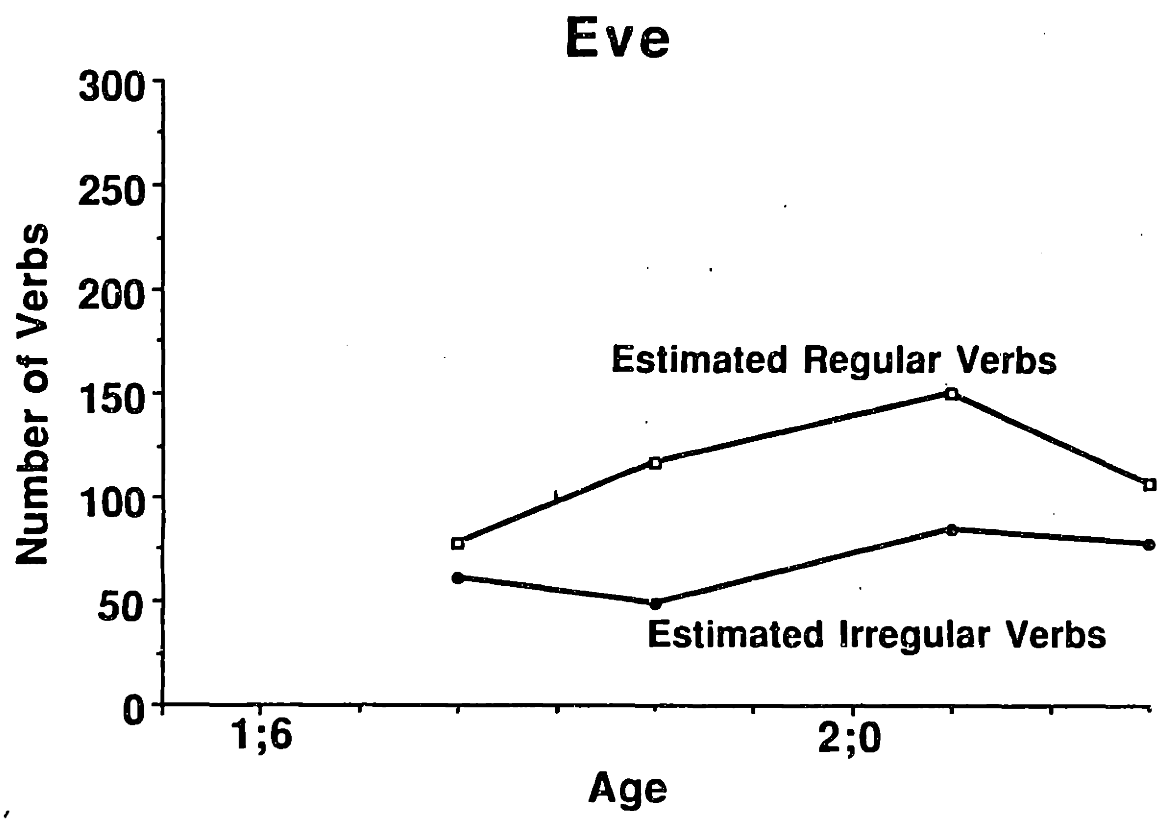


Figure 30. Proportion of Eve's Jackknife-estimated verb vocabulary that is regular, and her overregularization rate (subtracted from 100%).

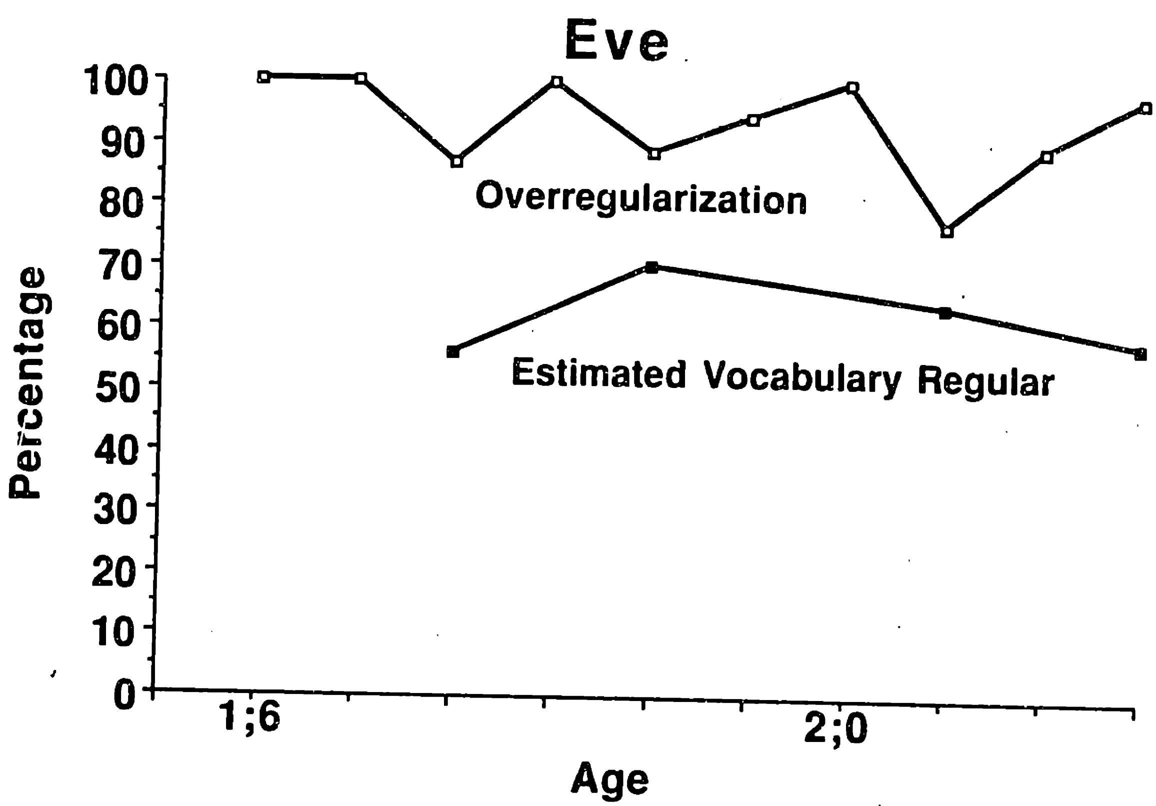


Figure 31. Sarah's Jackknife-estimated regular and irregular verb vocabulary.

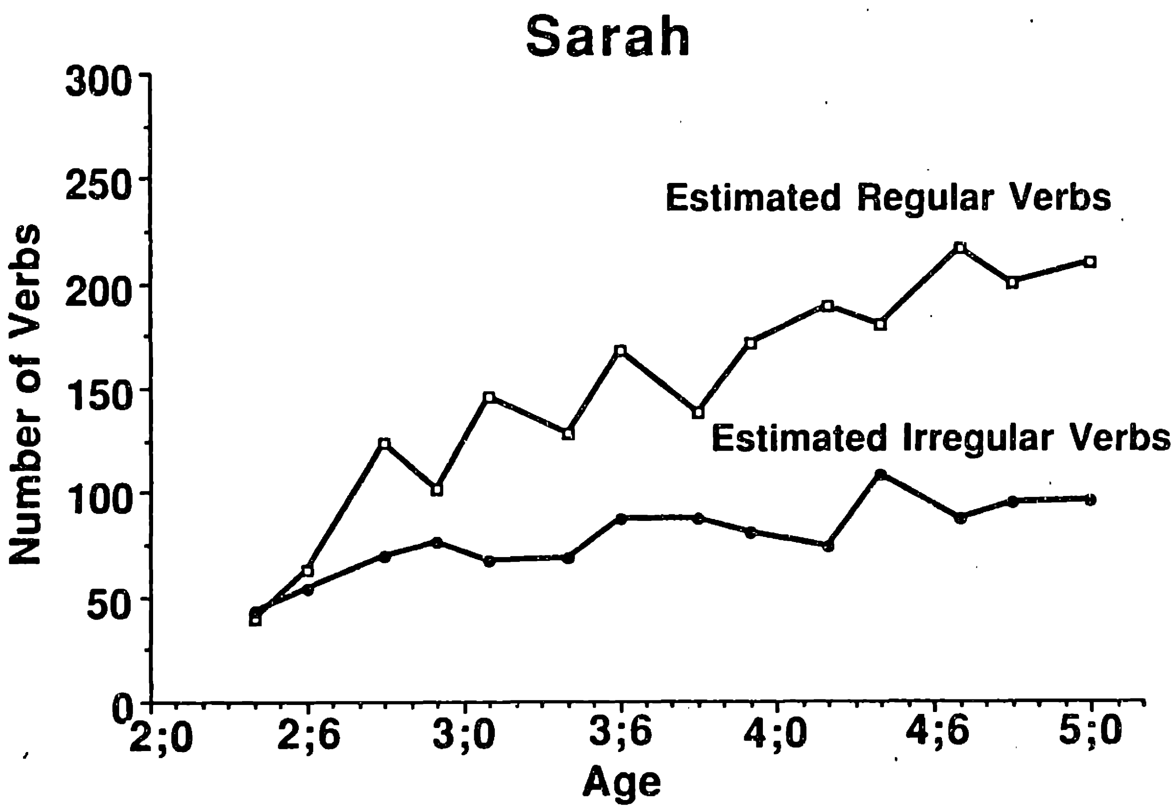


Figure 32. Proportion of Sarah's Jackknife-estimated verb vocabulary that is regular, and her overregularization rate (subtracted from 100%).

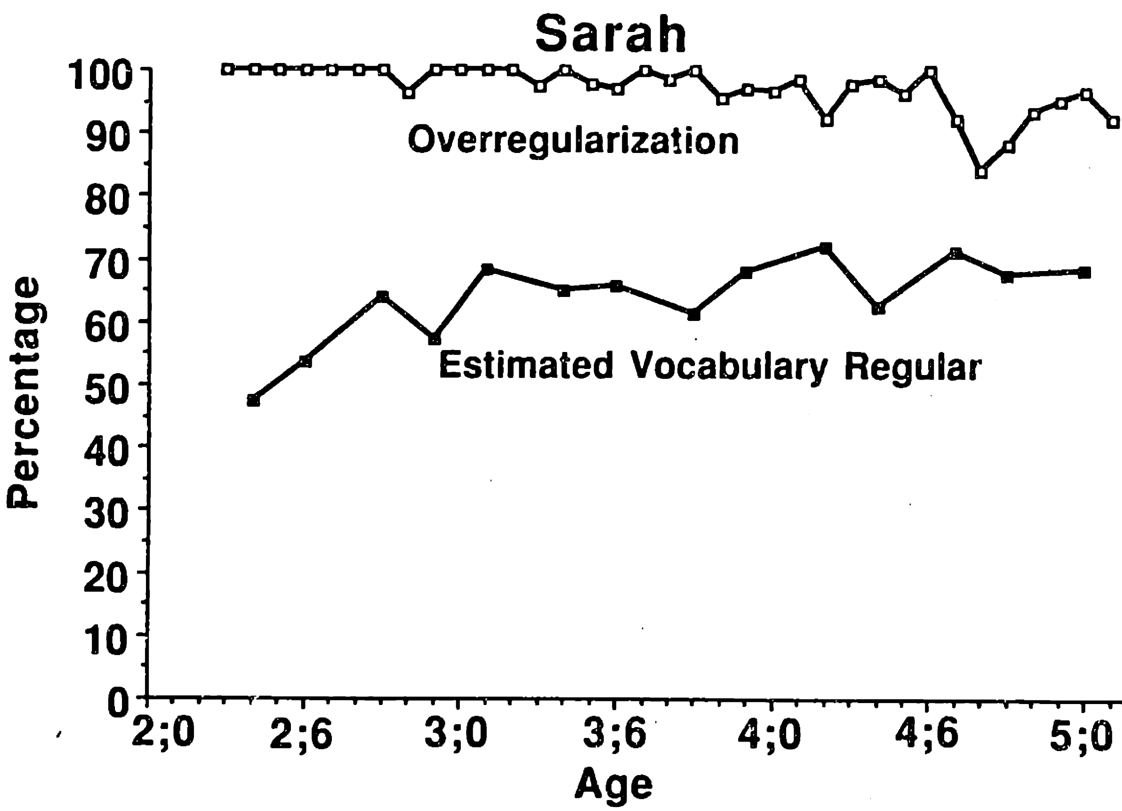


Figure 33. Proportion of Adam's irregular verbs in obligatory past tense contexts that were correctly marked for tense. Overregularizations do not enter into these data, but are shown in the curve at the top (subtracted from 100%).

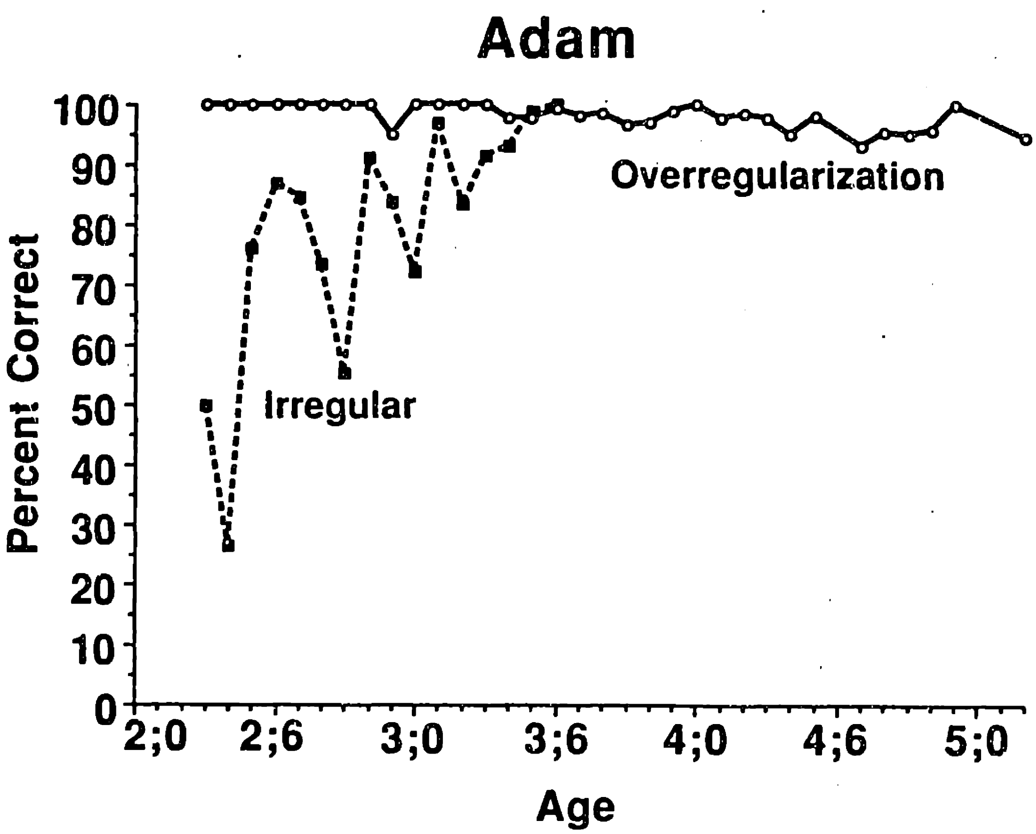


Figure 34. Proportion of Adam's regular verbs in obligatory past tense contexts that were correctly marked for tense, and his overregularization rate (subtracted from 100%).

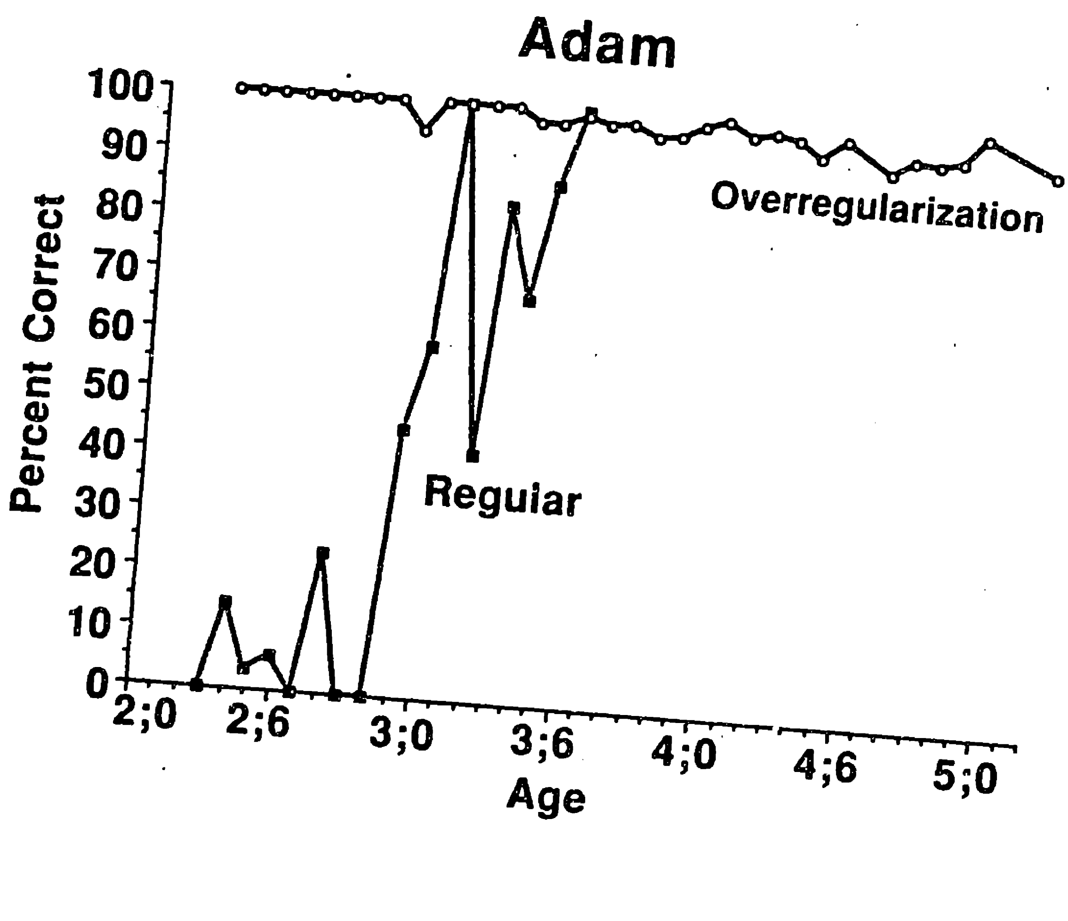


Figure 35. Proportion of Eve's irregular verbs in obligatory past tense contexts that were correctly marked for tense. Overregularizations do not enter into these data, but are shown in the curve at the top (subtracted from 100%).

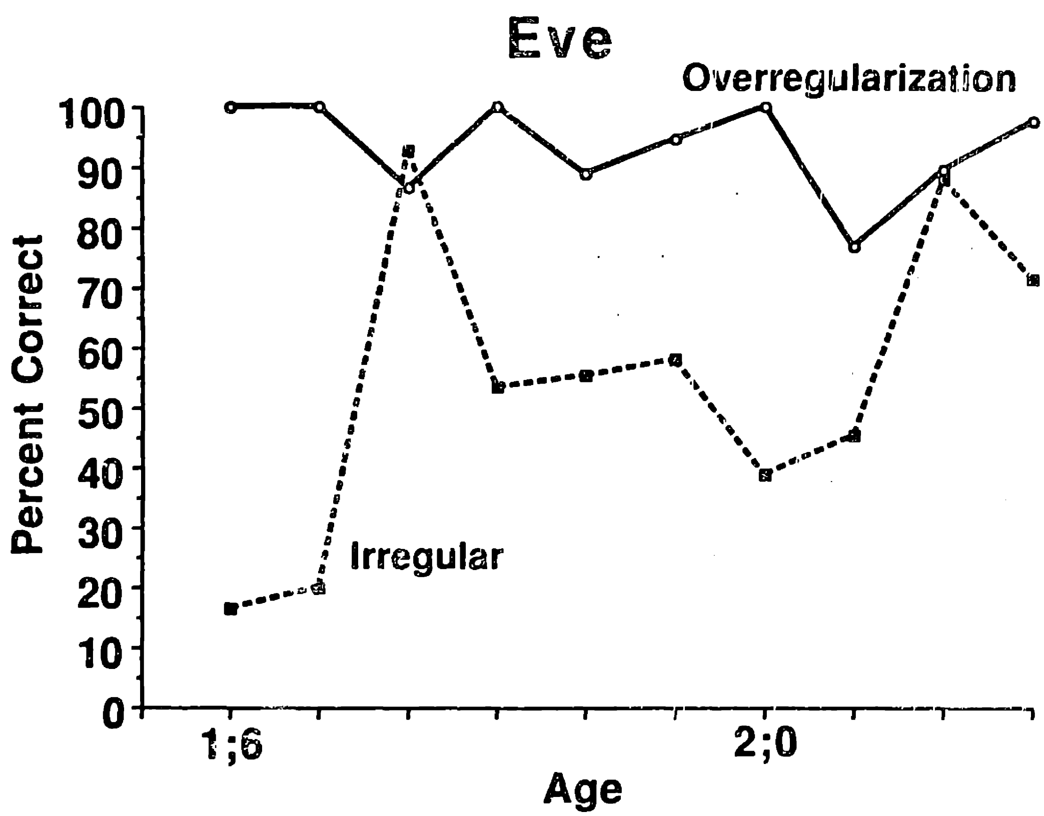


Figure 36. Proportion of Eve's regular verbs in obligatory past tense contexts that were correctly marked for tense, and her overregularization rate (subtracted from 100%).

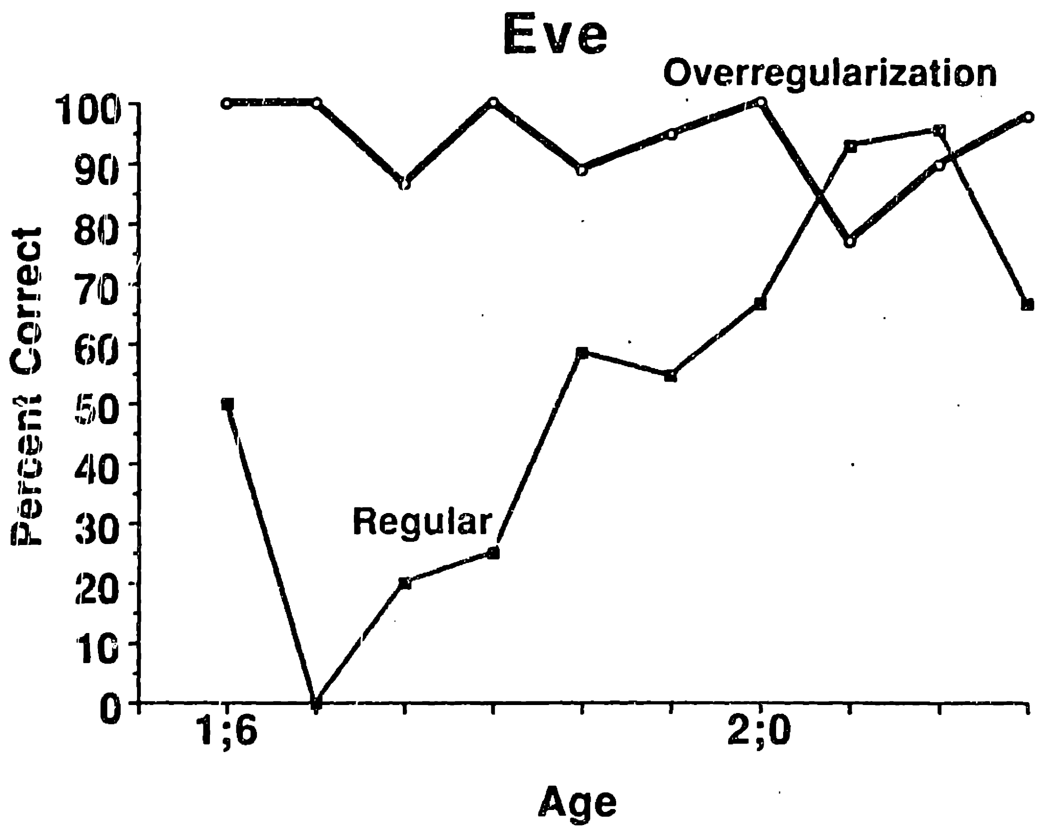


Figure 37. Proportion of Sarah's irregular verbs in obligatory past tense contexts that were correctly marked for tense. Overregularizations do not enter into these data, but are shown in the curve at the top (subtracted from 100%).

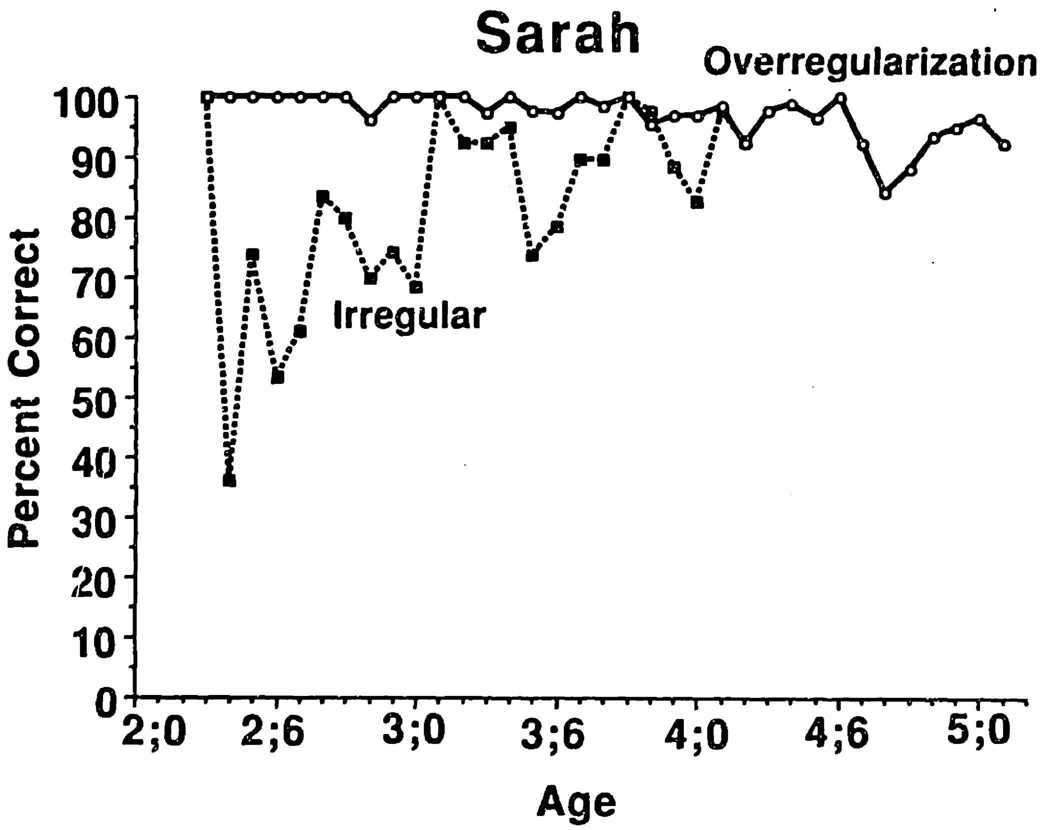


Figure 38. Proportion of Sarah's regular verbs in obligatory past tense contexts that were correctly marked for tense, and her overregularization rate (subtracted from 100%).

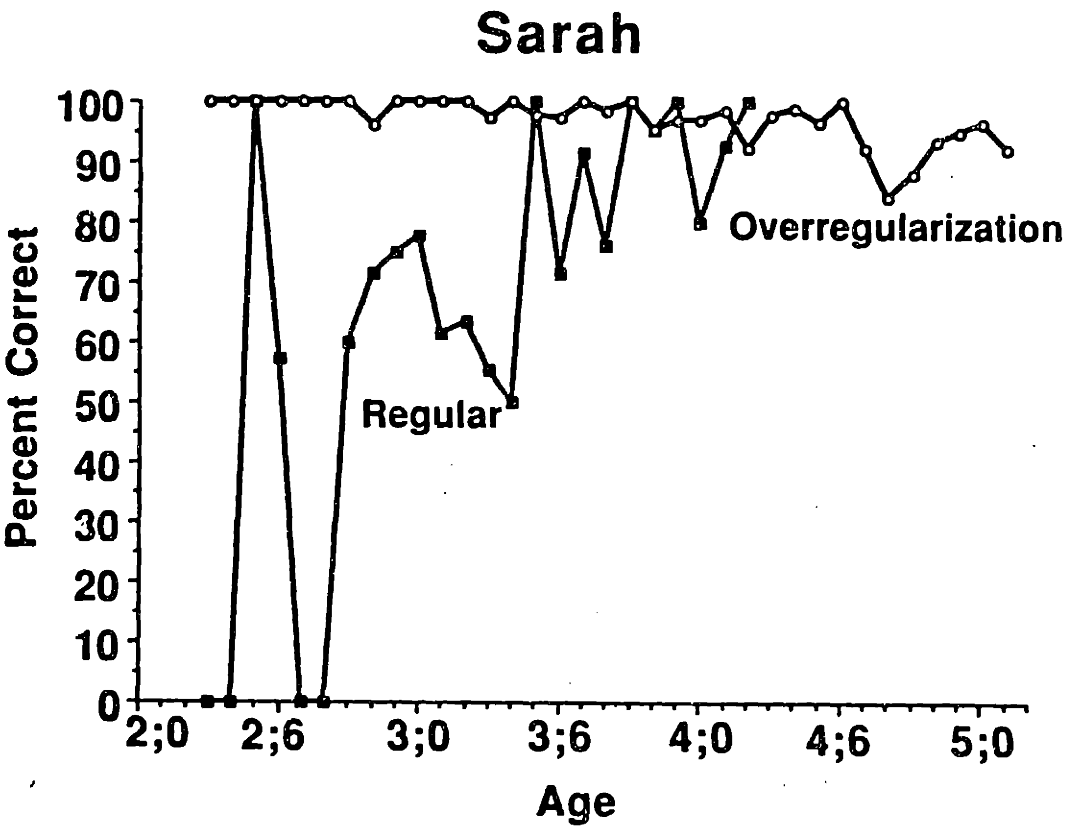


Figure 39. Proportion of Abe's irregular verbs in obligatory past tense contexts that were correctly marked for tense. Overregularizations do not enter into these data, but are shown in the curve at the top (subtracted from 100%).

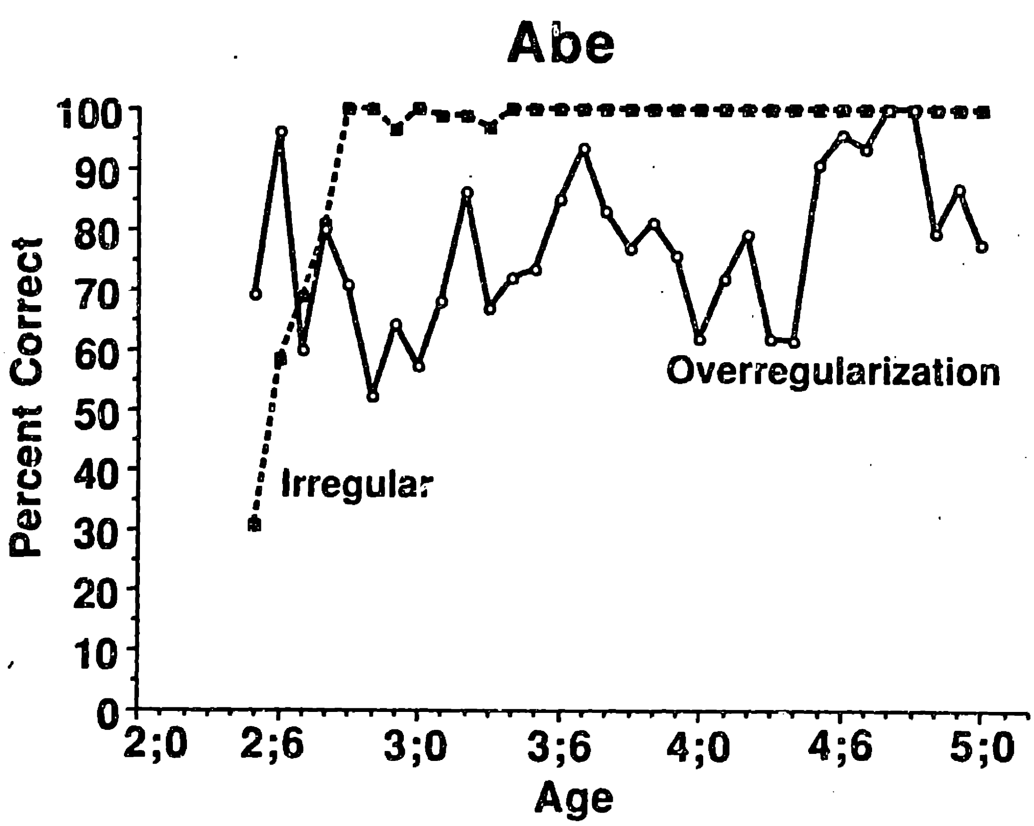
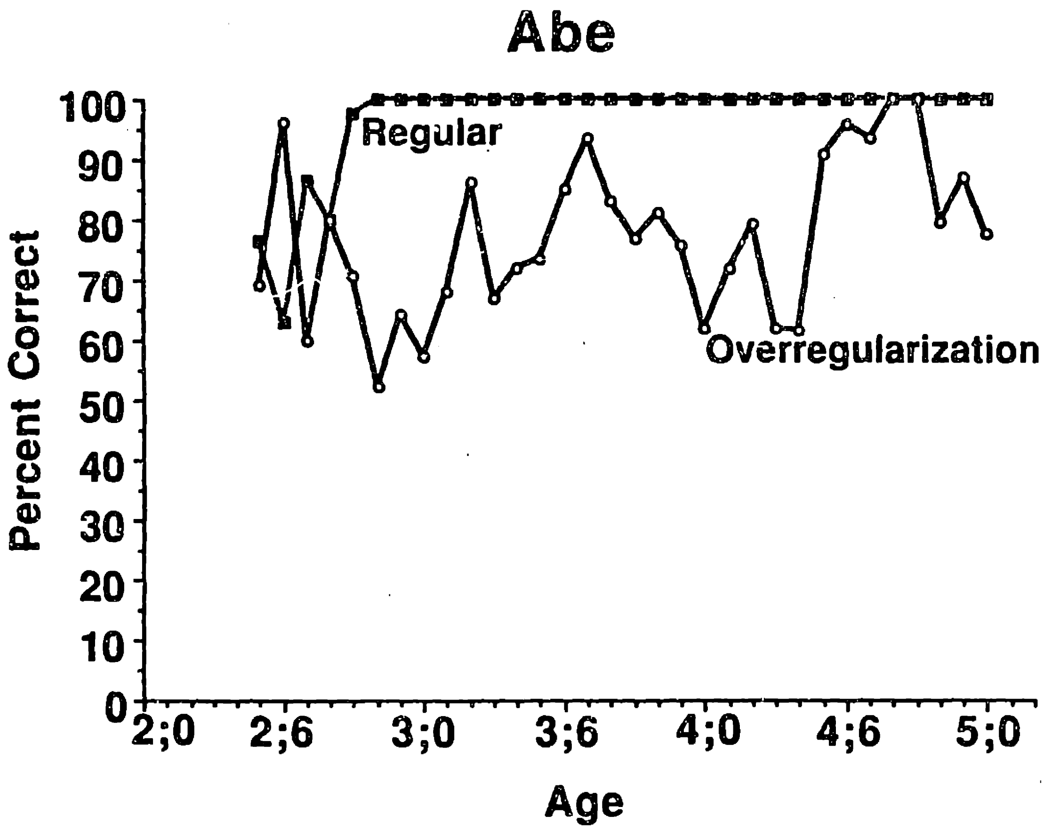


Figure 40. Proportion of Abe's regular verbs in obligatory past tense contexts that were correctly marked for tense, and his overregularization rate (subtracted from 100%).



PART III

Children's overregularization of English plurals: a quantitative analysis

Introduction

Past tense overregularizations, such as *comed* and *goed* are among the most widely discussed errors that children make as they learn language. The most common explanation of these errors is that children create them through the overapplication of a rule (e.g., to form the past tense, add *-ed*) to irregular verbs. Indeed, textbooks in developmental psychology often point to overregularization as the paradigmatic example of children's rule use. Children are said to follow a U- shaped developmental sequence in which they correctly inflect their first past tense forms, then go through a stage of overregularization, and finally unlearn these overregularizations and reach the adult stage. This U-shaped sequence has been used as a metaphor for psychological processes ranging from solving balance-beam problems to learning computer-user interfaces. (Karmiloff-Smith and Inhelder, 1974/1975; Grudin & Norman, 1991). Moreover, since children stop making these errors without explicit correction from their parents, overregularizations pose an intriguing learning puzzle (Brown & Hanlon, 1970; Baker, 1979; Marcus, 1993). Finally, these errors also occupy a central role in the recent debate between connectionist and symbolic models of language and cognition. (e.g., Rumelhart & McClelland, 1986; Pinker & Prince, 1988; MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991; Marcus, Pinker, Ullman, Hollander, Rosen, and Xu, 1992).

Until recently, however, the basic quantitative facts about past tense overregularization were unknown. In a recent *Monograph of the Society for Research in Child Development,* Marcus, et al. (1992) analyzed 11,521 past tense utterances from the spontaneous speech of 83 children. They discovered that -- contrary to popular opinion reported everywhere from *Newsweek* to textbooks to articles in the primary literature -- there is no stage in which children completely replace correct forms with overregularizations. Instead, children overregularize a small minority of the time, only about 4% of their opportunities. Furthermore, Marcus et al., confirmed quantitatively for the first time one aspect of the U-shaped developmental sequence: a period of correct performance does indeed precede the earliest overregularizations. To account for these data, Marcus et al. proposed a simple model, the *blocking-and-retrieval-failure* model -- Regular forms are formed by a default rule, "add -ed to form the past tense" (Berko's (1958) "Wug" experiment showed that children use this rule with novel verbs. See also Kim, Marcus, Hollander, and Pinker, 1991). Irregular forms are retrieved from memory and block the application of the default regular rule (add *-ed)*; if a child fails to retrieve the irregular form the regular rule applies by default and the child produces an overregularization. (See MacWhinney, 1978 for a similar model.)

Are plural overregularizations (e.g. *foots, mans*) produced by similar mechanisms as past tense overregularizations? Comparing noun plural overregularization with past tense overregularization allows a test of the replicability of Marcus et al.'s results. Moreover, comparison of developmental patterns of noun plural and past tense overregularization within individual children allows us to test whether overregularization is the result of an across-the-board grammatical change in the child's grammatical system or morpheme-by-morpheme development of morphology.

Verbs and nouns differ in many respects. For instance, Gentner (1978, 1982) has argued that nouns have "denser" representations than verbs; nouns are more stable and redundantly defined, while verbs are relational and more likely to change their meanings in context. Such semantic differences could conceivably result in differences between noun plural and verb past tense overregularization. On the other hand, if overregularizations are the result of a common morphological component that operates independently of the semantics input to it, then the developmental sequence of the two systems should be similar. According to the blocking-and-retrieval-failure model, plural overregularizations should be produced by a default rule (add *-s*) which applies whenever children cannot retrieve appropriate irregular forms. This model should predict that the time course of development of noun plurals should qualitatively resemble that of the acquisition of the past tense. For example, Marcus et al. showed that children use irregulars correctly before the first overregularization; if the acquisition of noun plurals is paced by similar mechanisms, then children should use some correct noun plurals before their first plural overregularizations.

Connectionist models

Rumelhart and McClelland (1986) showed that a connectionist model consisting of a single associative network, which contained no explicit rules of the sort "add *-ed*", was able to learn the past tense, while displaying a U-shaped sequence of development. The success of this model thus called into question the existence of psychologically real symbolic rules. While Rumelhart & McClelland's specific model has since been subjected to extensive criticism (e.g., Lachter & Bever, 1988; Pinker and Prince, 1988; Egedi & Sproat, 1991; Marcus et al., 1992; Marcus, Brinkmann, Clahsen, Wiese, Woest, and Pinker, 1993; Prasada & Pinker, 1993), much of which has been acknowledged by later connectionist modelers (e.g., MacWhinney & Leinbach, 1991; Plunkett and Marchman, 1991), several new connectionist models have been proposed.

Common to these connectionist models has been an alternative account of overregularization in which symbolic rules are eliminated in favor of a uniform associative process applies equally to regular and irregular words. In these models, overregularizations result because irregular forms are attracted to regular forms. For example, the overregularization *thinked* might occur by analogy to *blinked*. Thus, other things being equal, the more regular forms there are as compared to irregular forms, the more these models will overregularize. (See Plunkett & Marchman, 1991; Marcus et al., 1992, 1993; Prasada & Pinker, 1993 for further discussion.)

The study of noun plural acquisition offers a way to tease apart connectionist and rule-based models. Irregular noun plurals make up only 0.4% of noun types and 5% of noun tokens in contrast to irregular verbs which make up 14% of verb types and 60% of verb tokens (Marcus, et al., 1993). Thus, *ceteris paribus*, connectionist models would predict a greater rate of noun plural overregularization than verb past tense overregularization, since regulars vastly outnumber irregulars, both in terms of types and tokens. In contrast, the blocking-plus-retrieval-failure model would predict that a change in vocabulary balance would have no effect.

An even stronger difference between Rumelhart and McClelland's single associative network model and Marcus et al's blocking-and-retrieval-failure model is in the way these models account for the U-shaped sequence of development. According to Marcus et al, children's early stage of correct performance comes from using memorized forms; overregularization begins only when the child has mastery of the regular default rule. This model predicts that noun plurals should also have an initial period of correct use, before the child has induced the default plural rule, followed later by a period of overregularization after the child has induced the plural rule. In contrast, in Rumelhart & McClelland's (1986) model, the initial stage of the U-shaped curve (early correct irregulars) derives from the high percentage of vocabulary which is irregular in the early stages of training.[45] In their model, the first stage of training corresponds with the child's initial period of correct irregular use; at this point 80% of the input is irregular. The model begins to overregularize

---

[45]No other connectionist model displays this sort of U-shaped development. MacWhinney & Leinbach's (1991) model does not follow a U-shaped sequence of development. Plunkett and Marchman (1991) present several learning curves but none show an early period of 100% correct; see Marcus, et al., 1992 for further discussion.

when the input vocabulary becomes dominated by regular, rather than irregular, past tense forms. Because irregular noun plurals are so rare, there is unlikely to ever be a stage in which irregular plurals dominate regular plurals; hence the Rumelhart and McClelland model would probably overregularize even its earliest plurals.

The case of Abe

Maratsos (1993) has questioned whether children's past tense overregularization rates are as low as found by Marcus et al. His arguments center around a single child, Abe, who has an overregularization rate (0.24) that is substantially higher than the mean of the other 82 children that Marcus et al. studied (0.04). Marcus et al. themselves suggested several explanations for why Abe's overregularization rate is unusually high. One possibility is that Abe marks the regular past tense more reliably than other children (Marcus, et al., 1992; p. 112): if overregularizations are caused by the application of a default rule, then a child who applies the default rule more reliably should have a greater rate of overregularization. Marcus et al. (pp. 62-65) further suggest that some of the difference between Abe and the other children that they studied might be artifactual. Unlike all of the other children they studied, Abe was the son of a psycholinguist, Stan Kuczaj, who was specifically studying *past tense* ° overregularization errors. While Kuczaj does not detail the criteria by which Abe's transcribed speech was selected, he does mention that in a sample of 14 children used in his dissertation, Kuczaj asked parents to encourage their children to talk about specific events that took place in the past and about specific hypothetical events (both of which are marked, in English, by the past tense). Such elicitations could increase overregularization rates in several ways. First, elicitations may cause children to use verbs that they are unfamiliar with, and might otherwise avoid; such verbs are likely to be lower in frequency, and low frequency verbs more likely to be overregularized (Marcus, et al., 1992) . Second, some elictors contain the stem (e.g., *What did you eat?)* and hearing the stem might prime the overregularized form (e.g. *eat+ed*).

Third, elicitations resemble experimental situations, which tend to produce higher estimates of overregularization rates (Marcus, et. al, 1992).

Examination of noun plural overregularization can cast further light on this issue. If Abe's high rate of overregularization is due to some constitutional factor, as Maratsos appears to suggest, then one might expect Abe's *plural* overregularization rate to be higher than the plural overregularization rate of other children. If, on the other hand, Abe's high rate of past tense overregularization is partly artifactual, related in some way to the fact that his father studied him specifically to examine *verb past tense* overregularizations, then his rate of noun plural overregularization (which was not of interest to Kuczaj) should, *ceteris paribus*, simply be average.

Summary

The basic quantitative facts about noun plural overregularizations have not been documented. Although Brown (1973) and Mervis & Johnson (1991) provide examples of overregularized forms, neither provide numbers of correct forms, thus making it impossible to estimate rates of overregularization. Therefore, an analysis of plural overregularizations in spontaneous speech was conducted.

Method

Subjects

In order to allow direct comparison with previous results from verb overregularization, the 10 children studied by Marcus et al. that had their own individual databases were selected. Computerized transcripts containing conversations between these 10 children and their parents (and others), taken from the CHILDES database (MacWhinney and Snow 1985, 1990), were analyzed. Table 1 lists the children. All spoke standard American English; one came from a working class background, the remaining 9 from middle class backgrounds; half were boys, half girls.

Insert Table 1 About Here

Procedure

First, a list of 25 irregular noun plurals was compiled, using Quirk, Greenbaum, Leech and Svartik's (1985) *A comprehensive grammar of the English language* . These included one suppletive (*person-people*), two *-en* plurals (*child-children, ox-oxen)*, 6 vowel-change plurals (*foot- feet, goose-geese, louse-lice, mouse-mice, tooth-teeth* , *woman-women*), 13 voicing changes (*calves, elves, halves, knives, leaves, lives, loaves, selves, sheaves, shelves, thieves, wives, wolves*), 3 no-change plurals (*cod, sheep, deer*). Using the UNIX utility grep, all compounds (nouns or pronouns) found in the transcripts containing irregular plural heads (*barefeet, firemen, gentlemen, milkmen, ourselves, policemen, reindeer themselves, theirselves, workmen, yourselves*) were added to the list. Possible overregularized versions of these words, formed by adding the plural suffix -s to the stem (e.g., *mans*) or plural form (e.g., *mens*), were then added to the list of correct plurals in order to form a master list of search words. (Many, though not all, of children's overregularizations and other novel forms are appended by the CHILDES notations @n ; therefore, to insure all overregularizations were found, all forms appended with @n were also searched for.) Using the CHILDES program *kwal,* all uses of these search words were extracted from the transcripts. The resulting output was then hand-checked to evaluate whether the no change words were used as plurals or singulars and to eliminate forms not used as nouns (e.g., *lives* used as a verb) and plurals of names (which do not take irregular plurals, e.g., *All the grad students are supported by Fairchilds/*Fairchildren*). No-change plurals, were judged as plural only where indicated by the context (as determined by the author), and otherwise excluded. Because the focus of this paper is on *how* children mark plurals, rather than *whether* they decide to mark plural, all cases of marked plurals were included, hence the utterances "a foots" or "a feet" were both counted. The following were excluded: Irregular plurals such as *fish-fish* which allow an optional regular form (e.g., *fishes*), Greek and Latinate irregular plurals (e.g., *bacterium-bacteria*), humorous plurals (e.g. *Vax-Vaxen*), and the plural houses (/howzes/ vs. /houses/) which orthographically cannot be distinguished from the overregularized form. (Irregularizations, e.g., the hypothetical *house-hice*, were not searched for; in the case of verbs, irregularizations are an order of magnitude less frequent than overregularizations; see Xu & Pinker, 1992).

Following Marcus et al. (1992), a child's overregularization rate was calculated as the proportion of tokens of irregular past tense forms that are overregularizations:

$$\frac{\#\text{overregularization tokens}}{\#\text{overregularization tokens} + \#\text{correct irregular plural tokens}}$$

Overregularization tokens included both errors of the form stem + *s* (e.g., mans) and plural + *s* (e.g., mens). This equation measures *how* a child marks the plural given that the child marks the plural, rather than *whether* the child marks the plural. (See Marcus et al., 1992 for further discussion.)

Results

Insert Table 2 About Here

Table 2 gives each child's rate of overregularization. Overall rates of noun overregularization were low, replicating the low rates of verb overregularization found by Marcus et al. The mean rate of noun plural overregularization for these ten children was 8.5%, the median was 8.3% with rates ranging from 0% to 21.7%. Thus children systematically prefer the correct plurals to overregularized plurals. (One might ask whether the inclusion of voicing changes might have artificially lowered the rate of overregularization. However, children's rates of overregularization are not significantly different if the voicing changes are excluded, $t = 1.03$, $p = .33$, means 8.5% (including voicing changes) vs. 8.2% (excluding voicing changes). Henceforth, voicing changes are included in all analyses. (One might argue, since the three children with the smallest samples have the lowest overregularization rates, that the low overregularization rate is partly an artifact of small sample sizes, but note that Nathan has the second sparsest sample but the third highest overregularization rate, while Abe and Adam have the largest samples but average-to-low overregularization rates.)

These low overall rates of plural overregularization do not appear to be artifacts of averaging over time, but rather appear to be relatively consistent throughout development, as shown in Figures 1-4 (next to the symbol plotted for each month is the number of tokens for that month). Aside from a few months with very small sample sizes, overregularization rates are near 0. In all months with at least 20 plural tokens, the rate of overregularization is less than 20%.

Insert Figures 1, 2, 3, & 4 About Here

U-shaped development

Figures 1-4 also show that children's first plural overregularizations are preceded by a period of correct plural use, confirming one aspect of the U-shaped curve, as predicted by the blocking-and-retrieval model. (Marcus, et al., 1992) Of the ten children studied, only 7 overregularized at least once. All seven children used some number of correct plural before their first overregularization. To rule out the possibility that this is a statistical artifact resulting from the low overall rate of overregularization, the probability, for each child, that s/he could have had N correct forms given an overregularization rate P was calculated by the formula (1-p)^N. These probabilities are listed in Table 3. By treating each child as a separate study and combining the results by a metaanalysis (see Marcus et al, 1992, p.42 for further discussion), $p < .0025$.

Insert Table 2 About Here

Thus even though regular plurals heavily dominate irregular plurals, children initially use irregular plurals correctly, and only later begin to overregularize. This poses problems for connectionist models in which the U-shaped curve is the result of an initial dominance of irregular forms, but is consistent with models such as the blocking-and-retrieval-failure model in which regular forms are produced by a rule

Frequency

For each child, the correlations between the Francis & Kucera frequencies for each noun with the child's overregularization rate for that noun were tested. Four of the 7 children who overregularized at least once showed negative correlations, but the remaining 3 children showed positive correlations. The mean of these correlations (after transforming *r*s to *z*s, and transforming the resulting mean *z* back to an *r*) is -.15, almost the same as reported in Marcus et al., 1992, for verbs; the mean correlation between a measure of input frequency and child plural overregularization rates was .04. None of the correlations (in either direction) were significant. However, given the small number of nouns for each test (ranging from 8 to 12), there is not enough statistical power to accept the null hypothesis without a large risk of Type II error. (To detect a medium sized effect (r = .3) with power = .8, with alpha = .05, 85 noun types would be necessary for each child.) Thus the results concerning frequency are equivocal; substantially larger samples are necessary.

Comparison with verb overregularization

The mean rate of noun plural overregularization for these ten children (8.5%), was not significantly different from the mean rate of verb past tense overregularization for the same ten children (7.3%), $t(9) = .32$, NS. The fact that, relative to irregular verb past tenses, irregular noun plurals are quite rare, appears not have affected the rate of overregularization. This poses a challenge to uniform associative models of the acquisition of morphology which tend to be quite sensitive to the balance of regular and irregular vocabulary (see Plunkett & Marchman, 1991; Marcus et al., 1992, 1993; Prasada & Pinker for discussion).

Individual children might show similar developmental patterns across nouns and verbs. However, there was no correlation between rates of plural and past tense overregularization across months for Abe (r = -.12), Adam (r = -.03), Eve (r = -.01), or Sarah, (r = .05). This suggests that overregularization is not the result of a global, qualitative change, but is consistent with an account in which overregularizations are the result of occasional retrieval errors. Finally, Table 2 (above) shows that Abe's rate of plural overregularization, 7.4% is below the average rate of 8.5%, suggesting that his high rate of verb past tense overregularization is not due to a qualitatively different grammatical system. Instead, this lends some support to Marcus et al.'s speculation that Abe's high rate of past tense overregularization may have been partly due to the manner in which the data were collected. In support of this claim, Marcus et al., analyzed transcripts from four children: Abe, and the well-studied Adam, Eve, and Sarah (Brown; 1973), and found that children's rates of overregularization were greater following parental utterances which expected a past tense response (e.g., *Tell me what you did, What did you [verb]?*, etc.) than parental utterances which did not expect a past tense utterance Abe (27% vs. 19%), Adam (9% vs. 2%), Eve (44% vs. 15%), Sarah (25% vs. 18%). Maratsos (1993), using a slightly different procedure, failed to replicate this "parental elicitor" finding. Whereas Marcus et al tested every (stem) overregularization in Abe's samples; Maratsos restricted himself to a tiny subset of the data; he only used data during Abe's peak period of overregularization (2;7 to 3;4, selected post-hoc by Maratsos) and only for those verbs which are overregularized at rates of between .20 and .80 inclusive.[46] For those verbs, in that time

[46]The limits of .20 to .80 were selected post-hoc in order to examine only those verbs which "oscillated." Maratsos imposed these limits because "another problem is that particular verbs may be both more likely to be probed, and independently more likely to be overregularized (even when not probed by Abe's father), which would also produce a non-causal positive correlation between parental elicitors and overregularization rates." But then the possibility that the parent repeatedly elicits past tense of verbs that the child has trouble with, hence artifactually increasing the overregularization rate, is excluded a priori by the methodology.

range, Maratsos fails to find any difference between Abe's overregularization rate following "Parental elicitors" and Abe's overregularization rate when no elicitor was given. Similarly, Maratsos studies three of Kuczaj's sample of 14 children (apparently the data for the other 11 are lost) and again finds no effect. Maratsos argues from this non-effect that the past tense elicitor effect is not real.

Unfortunately, Maratsos's study doesn't have nearly enough statistical power to accept the null hypothesis without risking a Type II error. According to tables from Cohen (1992), with power = .80, and alpha = .05, the minimum number of utterances needed per group to detect a "medium" effect is 87; for a "small effect", the minimum number of utterances per group increases to 785. But Maratsos's study of Abe includes just 62 tokens following past tense elicitors (the three smaller samples examined by Maratsos, V.Q., I.B., and K.M contained just 7, 44 and 9 tokens after Past Elicitors, respectively). Marcus et al., using much bigger samples (e.g., for Abe, 375 utterances following past tense elicitors) were able to reject the null hypothesis. Thus the fact that Maratsos failed to replicate the earlier findings is probably due to the fact his samples are too small. Conclusions Nouns plurals and verb past tenses are overregularized at similar overall low rates and are likely to be produced by similar mechanisms. Moreover, the balance of regular and irregular plurals appears not to affect the rate or time course of overregularization, challenging connectionist models which attribute overregularization to the attraction of irregulars to regulars. Instead, these data support the blocking-and- retrieval failure model of overregularization.

References

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

Berko, J. (1958). The child's learning of English morphology. *Word* 14: 150- 177.

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

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

Bybee, J. L. and D.I. Slobin (1982). Rules and schemas in the development and use of English past tense. *Language* 58: 265-289.

Clahsen, H., Rothweiler, M. Woest, A., and Marcus, G. F. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition,* 225-255.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Egedi, D. M. & Sproat, R. W. (1991) Connectionist networks and natural language morphology. Unpublished manuscript, Linguistics Research Department, AT&T Bell Laboratories, Murray Hill, NJ.

Francis, N. & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Gentner, D. (1978) What looks like a jiggy but acts like a zimbo? A study of early word meaning using artificial objects. *Papers and Reports On Child Language Development*, 15, 1-6.

Gentner, D. (1982) Why nouns are learned before verbs: Linguistic relativity vs. natural partitioning. In S. A. Kuczaj II, ed., *Language development: Syntax and semantics.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Grudin, J. & Norman, D. A. (1991) Language evolution and human-computer interaction. *Proceedings of the 1991 Annual Meeting of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Karmiloff-Smith, A. & Inhelder, B. (1974/75) If you want to get ahead, get a theory. *Cognition,* 3, 195-212. Kim, J. J., Marcus, G. F., Hollander, M. and S. Pinker (1991). Children's inflection is sensitive to morphological structure. *Papers and Reports on Child Language Development* 30: 39-46. Kuczaj, S. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, 16, 589-600. MacWhinney, B. (1978) Processing a first language: the acquisition of morphophonology. *Monographs of the Society for Research in Child Development*, 43.

MacWhinney, B. & Leinbach, J. (1991) Implementations are not conceptualizations: Revising the verb learning model. Cognition. 40, 121-157.

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

Maratsos, M. (1993) Artifactual overregularizations? In E. Clark (ed.) *The proceedings of the twenty-fourth annual Child Language Research Forum.* Stanford, CA: Center for the Study of Language and Information.

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition*, 46, 53-85.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., Woest, A., and Pinker, S. (1993). German inflection: The exception that proves the rule. MIT Center for Cognitive Science Occasional Paper #47, Cambridge, MA.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J. and F. Xu (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development.*, 57 (4, Serial No. 228).

Mervis, C. B. & Johnson, K. E. (1991) Acquisition of the plural morpheme: a case study. *Developmental psychology*, 27, 222-235.

Pinker, S., and Prince, A. (1988) On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition. *Cognition* 28: 73-193.

Pinker, S. and Prince, A. (1992). Regular and irregular morphology and the psychological status of rules of grammar. In L. A. Sutton, C. Johnson, & R. Shields (Eds.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on the Grammar of Event Structure*. Berkeley, CA: Berkeley Linguistics Society.

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.

Prasada, S. & Pinker, S. (1993) Similarity-based and rule-based generalizations in inflectional morphology. *Language and Cognitive Processes*, 8.

Quirk, R., Greenbaum, S., Leech, G., & Svartik, J. (1985) *A comprehensive grammar of the English language*. New York: Longman.

Rumelhart, D. and J. McClelland (1986). On learning the past tenses of English verbs. Implicit rules or parallel distributed processing? In: J. McClelland, D. Rumelhart and the PDP research group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*. MIT Press: Cambridge, MA.

Xu, F. & Pinker, S. (1992) Weird past tense forms. Seventeenth annual Boston University Conference on Language Development. Boston, MA.

)

Appendix: Children's rates of overregularization for individual verbs

	Abe	Adam	Allison	April	Eve	Naomi	Nat	Nathan	Peter	Sarah
barefeet	0.00	NA	NA	NA	NA	NA	NA	NA	0.00	NA
calves	NA	NA	NA	NA	NA	0.00	NA	NA	NA	NA
children	0.25	0.00	0	NA	0.00	0.00	0.00	0.00	0.11	0.05
deer	0.00	NA	NA	NA	NA	NA	NA	0.00	NA	NA
elves	0.57	NA	NA	NA	NA	NA	NA	0.00	NA	NA
feet	0.00	0.08	0	0	0.07	0.08	NA	0.00	0.16	0.05
firemen	0.50	0.44	NA	NA	NA	NA	NA	NA	NA	NA
geese	0.67	0.00	NA	NA	NA	0.00	NA	NA	NA	NA
gentlemen	0.00	0.00	NA	NA	NA	NA	NA	NA	NA	0.00
halves	NA	NA	NA	NA	NA	0.00	NA	NA	NA	NA
knives	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA
leaves	0.00	0.00	NA	NA	0.00	0.00	NA	0.67	0.00	0.33
lice	NA	NA	NA	NA	NA	0.00	NA	NA	NA	NA
men	0.10	0.14	NA	NA	0.29	0.46	NA	0.00	0.50	NA
mice	0.14	NA	NA	NA	NA	0.50	NA	0.00	0.00	0.00
milkmen	NA	0.00	NA	NA	NA	NA	NA	NA	NA	NA
ourselves	0.22	NA	NA	NA	NA	NA	NA	NA	NA	0.00
people	0.05	0.02	NA	0	0.08	0.09	0.00	0.50	0.23	0.00
policemen	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA
reindeer	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.86
selves	NA	0.33	NA	NA	NA	NA	NA	NA	NA	NA
sheep	0.14	0.00	0	NA	0.50	NA	NA	NA	0.43	NA
teeth	0.05	0.00	NA	0	0.00	0.00	NA	0.00	0.00	0.00
theirselves	NA	NA	NA	NA	NA	0.00	NA	NA	NA	NA
themselves	0.00	NA	NA	NA	NA	NA	NA	NA	0.00	NA
women	0.00	NA	NA	NA	NA	NA	NA	NA	NA	0.00
workmen	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA
yourselves	NA	NA	NA	NA	NA	NA	NA	NA	0.00	NA

Table 1
Children Studied

Child	Age	Source	Total Samples	Sampling Frequency
Abe	2;6-5;0	Kuczaj (1977)	210	weekly
Adam	2;3-5;2	Brown (1973)	55	2-3/month
Allison	1;5-2;10	Bloom (1973)	6	occasional
April	1;10-2;11	Higginson (1985)	6	occasional
Eve	1;6-2;3	Brown (1973)	20	2-3/month
Naomi	1;3-4;9	Sachs (1983)	93	weekly to monthly
Nat	2;8	Bohannon (1977)	21	within 1 month
Nathaniel	2;3-3;9	Snow	30	weekly
Peter	1;3-3;1	Bloom (1973)	20	monthly
Sarah	2;3-5;1	Brown (1973)	139	weekly

TABLE 2
Overregularization Rates for individual children

Child	Correct	Stem+s	Plural+s	Total	Overreg. Rate
Abe	284	22	1	307	0.075
Adam	243	9	6	258	0.058
Allison	6	0	0	6	0
April	5	0	0	5	0
Eve	68	3	6	77	0.117
Naomi	81	8	6	95	0.147
Nat	5	0	0	5	0
Nathan	53	0	9	62	0.145
Peter	173	14	34	221	0.217
Sarah	109	10	1	120	0.092

Table 3
Correct irregulars preceding the sample with the first overregularization

Child	Age at 1st Overreg.	Sample #	Correct plurals before 1st overreg	OR Rate	Prob
Abe	2;6	13	5	0.0749186	0.677484
Adam	2;7	11	24	0.0581395	0.237505
Eve	1;10	10	13	0.116883	0.198718
Naomi	2;4	55	22	0.147368	0.0299748
Nathan	2;7	8	21	0.145161	0.0371168
Peter	2;6	14	14	0.217195	0.0324452
Sarah	3;1	43	22	0.0916667	0.120612

Figure 1: Percentage of Abe's irregular plurals that are correct (100% minus the overregularization rate.)

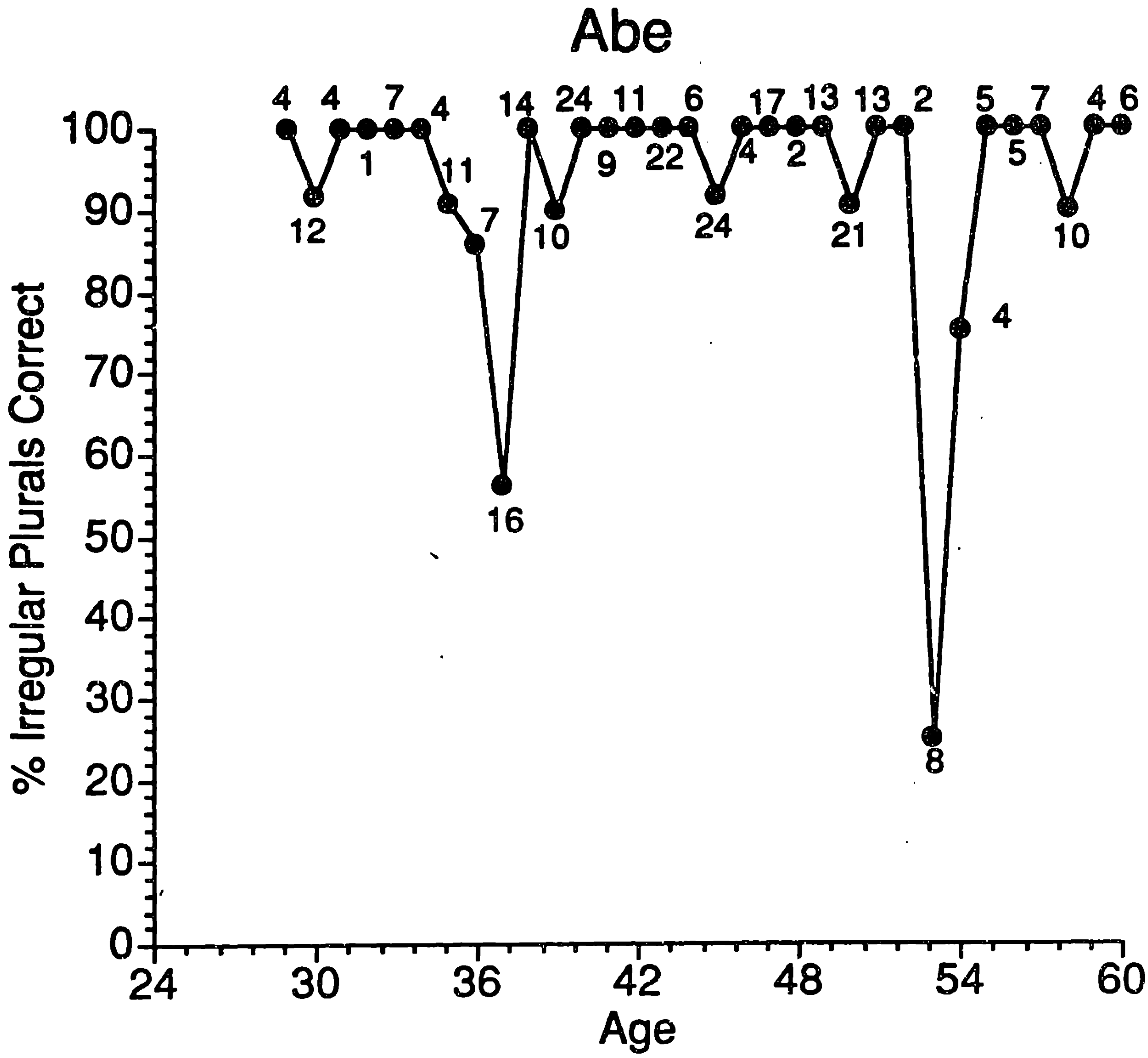


Figure 2: Percentage of Adam's irregular plurals that are correct (100% minus the overregularization rate.)

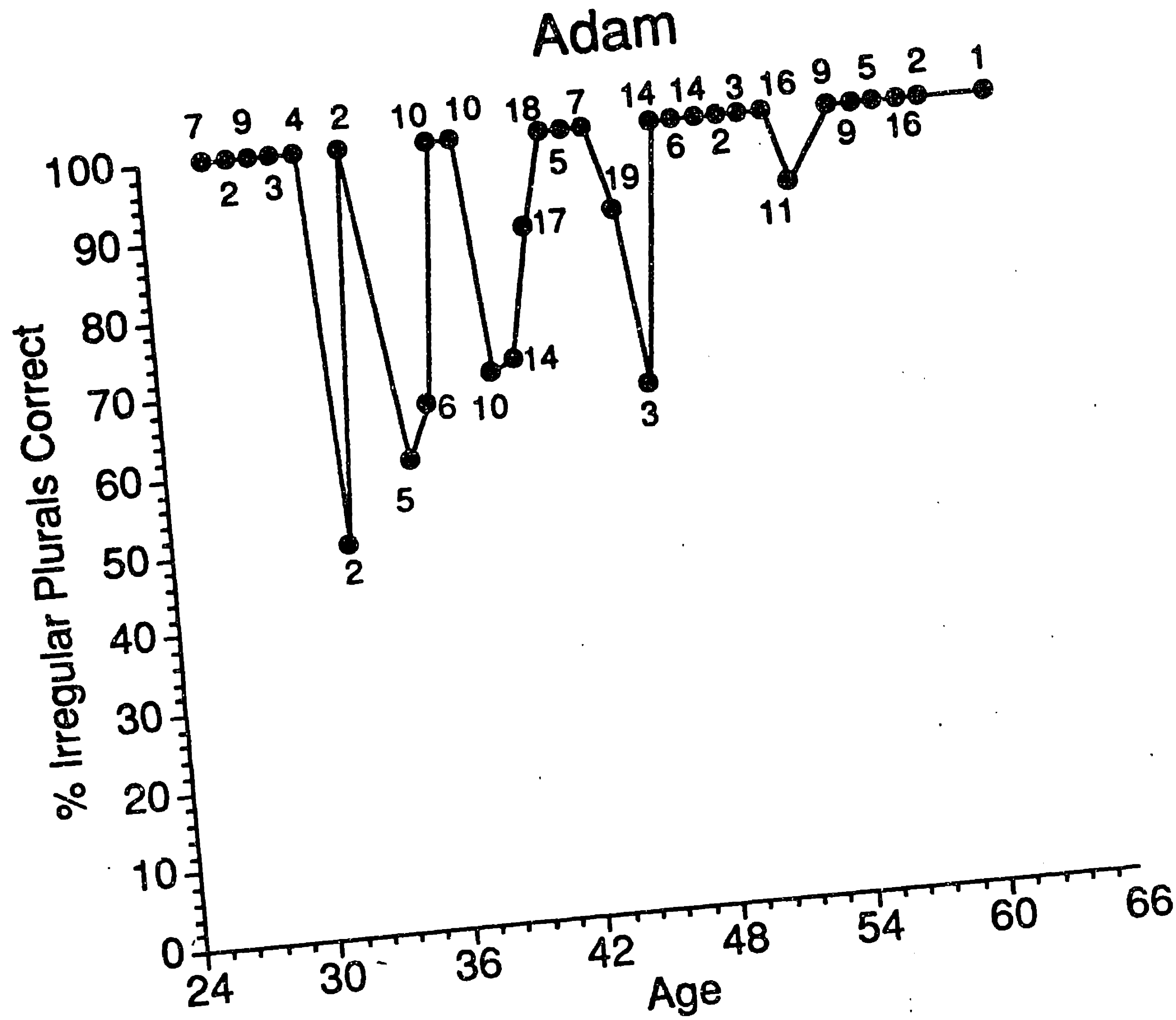


Figure 3: Percentage of Eve's irregular plurals that are correct (100% minus the overregularization rate.)

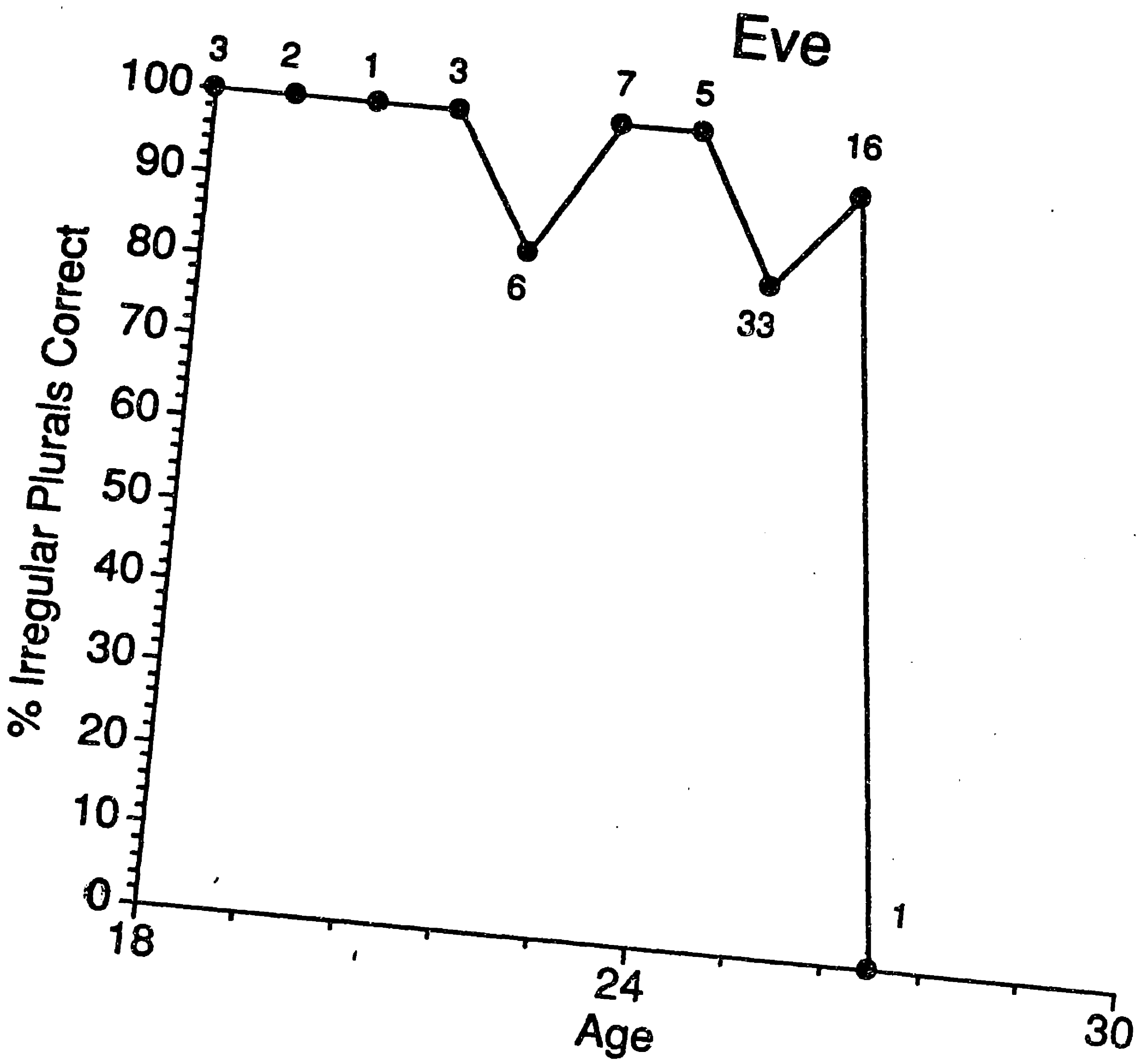
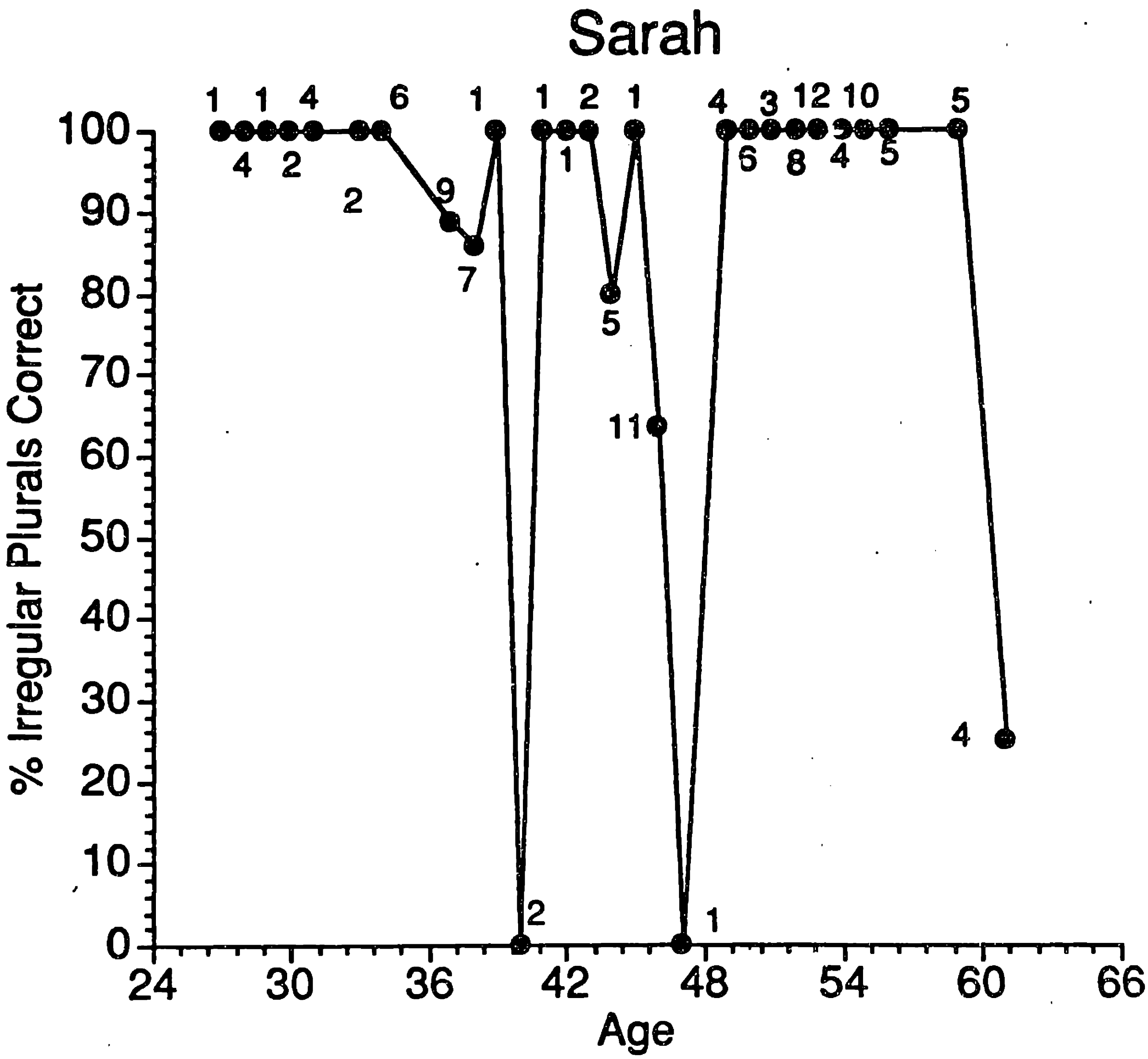


Figure 4: Percentage of Sarah's irregular plurals that are correct (100% minus the overregularization rate.)



PART IV

German Inflection: The Exception that Proves the Rule

Introduction

In 1980, the late Allen Newell singled out one discovery as the central contribution of cognitive science, comparable in importance to natural selection in evolution, the cell doctrine in biology, plate tectonics in geology, and the germ theory of disease (Newell, 1980; see also Newell & Simon, 1981; Fodor, 1968). According to the Physical Symbol System Hypothesis, intelligence (in humans or computers) is the product of mechanically-implemented rules manipulating symbolic representations or data structures. Human language has always been seen as a paradigm case of such a generative rule system. Chomsky (1957, 1965) pointed out that all normal humans can produce and understand an infinite number of novel sentences, including sentences whose words have no associations with familiar meanings or with each other, like *Colorless green ideas sleep furiously.* This ability can be explained on the theory that people have memorized a finite list of simple words in a mental dictionary, and combine them using a finite list of rules in a mental grammar. The rules specify the arrangement of abstract symbols like Noun, Verb, Noun Phrase, and Verb Phrase, rather than words directly. Thus the rules can generate an infinite number of discretely different sentences, by specifying, for example, that a Verb Phrase may contain a Verb Phrase (thereby generating *I think that he thinks that she thinks that ...*). They can also generate sentences with completely unfamiliar sequences of words, by specifying, for example, that any noun can combine with any verb, regardless of how familiar or sensible the particular noun and the particular verb are in combination (thereby generating *colorless green ideas* using the same rules as in *strapless black dresses.*)

Recently, however, one domain of language has been the subject of a debate over the empirical status of rules as mental, and ultimately neural, entities. English inflectional morphology at first seems like a transparent illustration of the use of rules and their contrast with memorized lexical items. Nearly all English verbs form their past tense in the same way, by adding the morpheme *-ed*.[47] The past tense of these verbs (e.g., *walk-walked*, *perambulate-perambulated*) is said to be "regular." People productively extend the regular pattern to new verbs, like *faxed* and *out-Gorbacheved* (what Boris did to Mikhail). Even preschool children, after hearing a novel verb like *rick* in the laboratory, easily create its regular past tense form *ricked* (Berko, 1958). The predictability and open-ended productivity of the regular pattern suggests that regular past tense forms are generated, when needed, by a mental rule: "to form the past tense of a verb, add the suffix *-ed*."

In contrast, English contains about 180 exceptional or "irregular" verbs that form their past tense in idiosyncratic ways, such as *ring-rang*, *sing-sang*, *go-went*, and *think-thought*. The idiosyncrasy and fixed number of irregular verbs suggests that they are memorized as a list of exceptions. The memory and rule components interact in a simple way: retrieval of an irregular form from memory blocks application of the rule, preventing *breaked*. Children occasionally make "overregularization" errors like *breaked*, presumably because they occasionally fail to retrieve the irregular form from memory and thus fail to block application of the rule (see Marcus, Pinker, Ullman, Hollander, Rosen, and Xu, 1992 for a review).

To understand the psychological issues at stake, it is crucial to distinguish two very different senses of the notion "rule" and its related concept "regular." One sense of "rule" is as a purely statistical or descriptive generalizations holding for most or all the phenomena they describe, like "planets follow elliptical trajectories." In familiar philosophical

---

[47]There are three pronunciations of this morpheme -- the [t], [d], and [ɨd] in *walked*, *jogged*, and *patted*, respectively -- but they represent a predictable phonological alternation that recurs elsewhere in the language. Hence they appear to be the product of a separate process of phonological adjustment applying to a single underlying morpheme, /d/; see Zwicky (1975) and Pinker & Prince (1988).

terminology, one would say that the planets "obey the rule" (i.e., merely conform to it) but do not "follow the rule" (i.e., actually consult it). Similarly, runners conform to the rule "All humans run marathons in more than two hours," but they certainly do not need to recall any such rule stored in their brain; they simply run as fast as they can and rule-conformity happens. This is not the sense of "rule" that Newell, Fodor, and Chomsky refer to. They posit actual mental operations that manipulate internally-represented symbols. When spellers actually think to themselves, "*i* before *e* except after *c*," and decide how to type a word accordingly, they are actually following or consulting a rule (though following rules need not, in general, be conscious). Such psychological rules need not be statistically general; they must apply when invoked, but may be invoked only rarely. Thus one version of the claim that the *-ed* past tense suffix in English is "regular" or "the rule" would just be a claim about vocabulary statistics, validated by simply opening a dictionary and counting. The more interesting version of the claim, and the one discussed in this paper, is that the rule generating the regular suffix is a mental operation. The problem is how to test that claim.

Rumelhart and McClelland (1986) have explicitly suggested that regularity in the English past tense system (and rules in general) may be a fact about the language but not about language users. They designed a connectionist pattern associator network in which inflection was computed not by rules, but by a network consisting of two connected layers of neuron-like units which corresponded to bits of phonological information in the stem and past tense form, respectively. After extensive training, the model picked up the statistical contingencies between stem sounds and past tense sounds, and thereby produced the correct past tense forms of 420 verbs, generalized to many new verbs, regular and irregular, produced overregularizations like *breaked* during intermediate stages of learning, and had more trouble, on average, with the kinds of irregular verbs that children have more trouble with. Rumelhart and McClelland concluded, "We have, we believe, provided a distinct alternative to the view that children learn the rules of English past-tense formation in any explicit sense. We have shown that a reasonable account of the acquisition of past tense can be provided without recourse to the notion of a 'rule' as anything more than a *description* of the language" (p. 267).

English inflection is probably not the fairest arena in which to test the psychological status of rules. Unlike phrases and sentences, tensed verbs are of a fixed length, display no recursion, and involve a combination whose second member is drawn from a list of three (*-s*, *-ing*, *-ed*). Thus they do not require the full resources of rule systems. Nonetheless, the English past tense system has recently been the subject of intense research on the psychological reality of rules, with, among others, Pinker & Prince (1988, 1992; Pinker, 1991), Marcus, et al. (1992), Lachter & Bever (1988), Clahsen, Rothweiler, Woest, & Marcus (1993), and Sproat (1992) adducing evidence that a rule is necessary, and Plunkett & Marchman (1990, 1991), MacWhinney & Leinbach (1991), Hare and Elman (1992), and Seidenberg & Daugherty (1992), among others, suggesting that a pattern associator is sufficient. The more general debate on symbolic versus connectionist cognitive architecture (e.g., Fodor & Pylyshyn, 1988; Smolensky, 1988; McClelland, 1988) often focuses on English inflection as a test case.

In this paper we first review the status of the debate: the aspects of the English past tense system that make a pattern associator plausible, the evidence that has been adduced for a mental rule, and connectionists' attempted counterexplanations for this evidence. Then we present two experiments that involving a new and more decisive kind of evidence: German inflection. To preview, we will suggest that the debate over a past tense rule was left open because of an accidental confound in English: the regular past tense is rule-governed both descriptively (it characterizes the vast majority of verbs) and psychologically (speakers can generalize the rule to any new word that bears the mental symbol "verb," regardless of its availability from memory). Connectionist pattern associators exploit this confound, because they are designed to generalize the most pervasive and frequent patterns. German breaks the confound: there is a participle suffix that is not regular or rule-governed in the descriptive sense (it does not apply to the vast majority of verbs) but is regular and rule-governed in the psychological sense (speakers generalize it to any new word that bears the mental symbol "verb," regardless of its availability from memory). A second experiment, involving a plural form, makes this point more strongly, as the suffix applies only to a tiny minority of nouns in German but it is generalized freely. This argues that a mental rule is necessary for regular inflection, and is not an epiphenomenon of the statistical preponderance of regular words and the ability of pattern associators to exploit it.

The Partial Structure of Irregular Morphology

If English inflection simply consisted of 180 completely idiosyncratic irregular past forms and a suffix that applied to all other verbs, there would be little to discuss; a rote-memorized list for the irregular verbs and a concatenation operation that applied elsewhere would be the simplest conceivable theory. The pattern-associator challenge is motivated by an important wrinkle in this picture: the irregulars appear to be somewhere in between a completely idiosyncratic fixed list and a product of a freely-applying set of rules.

The English irregular verbs display three kinds of partial patterning (see Pinker & Prince, 1988). First, with the exception of *go-went* and *be-was*, irregular past tense forms share much of their phonological form with their corresponding stems. For example, *ring* and *rang* differ only in their medial vowel. Second, there are far fewer kinds of changes from stem to past than there are irregular verbs. For example, the *i-a* alternation seen in *ring-rang* is also found in *sing-sang*, *spring-sprang*, *drink-drank*, *stink-stank*, and *spit-spat*. Third, irregular verbs displaying these recurring alternations are "hypersimilar" (Pinker & Prince, 1988): even though the only logical requirement for undergoing an *i-a* alternation is that the stem contain an *i*, the verbs that undergo this an alternation are similar in many other aspects. For example, the *i-a* verbs often have an initial consonant cluster containing *s*, a stop consonant, and a liquid consonant, and a final consonant cluster with velar and nasal features. These three redundancies suggest that the mental links between stem and past are not completely arbitrary, as a rote list of slot-pairs for paired-associate items would be.

Chomsky and Halle (1968) and Halle and Mohanan (1985) propose that irregular past tense forms are produced through the application of morphologically conditioned phonological rules. For instance, *rang* would be generated by taking *ring* and applying the rule of Lowering Ablaut, which specifies that *i* changes to *a*. These "minor rules" explain two out of the three kinds of irregular patterns. Stem-past similarity, such as between *ring* and *rang*, comes from the fact that the process creating the past tense form literally takes the stem as input and modifies it. If the rule simply changes *i* to *a*, then the rest of the stem comes through in the output untouched. Pair-pair similarity, such as between *ring-rang* and *sit-sat*, comes from the fact that a single set of rules is shared among a larger set of verbs; *ring* and *sit* share a single Lowering Ablaut rule.

The third kind of patterning, hypersimilarity, is not well handled by the minor-rule theories. The problem is that it is extremely difficult to characterize necessary and sufficient conditions for the application of minor rules. For instance, Lowering Ablaut applies to many irregular verbs: *ring, sing, spring, drink, sink, shrink, stink, swim, begin, sit*, and *spit.* But it fails to apply to many others, such as such as *bring, think, cling, fling, sling, sting, string, swing, wring, slink, stick, dig, win, spin, give*, and *bid.* Note that this problem would arise even if Lowering Ablaut were constructed so as to apply to the pattern common to the stems like *spring* that do undergo it, such as "*i* followed by a consonant cluster and preceding a velar nasal." Such a pattern would still result in irregular false alarms like *bring* and *string*, as well as regular ones like *wink* and *blink.* For this reason, Halle and Mohanan propose that each minor rule must specify the exact set of words it applies to: "The grammar will not contain a plethora of statements, such as 'the past tense of sing is sang, the past tense of bind is bound,' etc. Rather, it will contain a few rules, each of which determines the stem vowels of *a list of verbs specifically marked to undergo the rule in question.*" (p. 104, emphasis ours).

A problem for this proposal is that the list of verbs specifically marked to undergo the rule in question, while not characterizable by any necessary and sufficient pattern, display partial patterning (hypersimilarities) that are not accounted for by simply stipulating that they are members of a list. For example, among verbs undergoing the *i-a* rule there are far more with initial consonant clusters, and with final velar nasals than one would expect by chance. Moreover, there are other verbs that only partially fit this prototype, like *swim-swam*, whose final consonant is nasal but not velar, and *spit-spat*, whose final consonant is neither nasal nor velar. Other families of verbs undergoing a given irregular alternation also display these family resemblances (Bybee and Slobin, 1982; Bybee and Moder, 1983; Pinker & Prince, 1988). The dilemma, then, is that a template for verbs undergoing the rule yields many false alarms and misses, while a rote list of verbs undergoing the rule fails to explain the patterns of hypersimilarity or family resemblance among them.

A related problem is that people are sensitive to these hypersimilarities and sporadically use them to extend irregular alternations to new verbs, not on the list. Children occasionally make errors like *trick-truck*, *bring-brang*, and *wipe-wope* (Bybee and Slobin, 1982; Pinker & Prince, 1988; Xu and Pinker, 1992), generally with verbs that are highly similar to the corresponding irregular. Adult experimental subjects, when faced with inflecting a novel verb like *spling*, sometimes produce *splang* or *splung* (Bybee and Moder, 1983; Prasada and Pinker, 1993). This tendency directly reflects the degree of hypersimilarity: people are more likely to inflect *spling* as *splung* than they are to inflect *nist* as *nust*. Some English irregulars must have come into the language by this route: *caught*, *knelt*, *quit*, and *snuck* are relatively recent additions to English, and alternative forms like *heat-het* and *bring-brung* can be found in nonstandard dialects (Pinker & Prince, 1988). Existing irregulars that are low in frequency but match a prototype, like *slink-slunk* or *thrive-throve*, often feel neither exactly right nor clearly wrong to native speakers (Pinker & Prince, 1988; Ullman and Pinker, 1993).

One way out of this dilemma is to enrich the notion of memory. Say the memory structure for words to which the alternations are linked is not just a fixed list of atomic items, but also a list of the phonological patterns they contain, and that the links are not all-or-none but of graded strength. That is, *i-a* is not just joined to a representation of the words *ring* and *spring*, but to a representation of *-ing-* and *-ri-*, which are shared by *ring, drink, spring, bring*, and so on. Though the link between word and alternation ordinarily suffices to retrieve the correct alternation for the word and generate the correct past tense form, this is more likely when the link to the alternation is reinforced by links to to the words' various patterns. Conversely, when a word is not linked to an alternation but one of its phonological components is linked to an alternation (because other words containing that component are so linked), the word is prone to being irregularized. Thus lists of irregulars will be remembered more reliably when they are hypersimilar, and new, similar irregulars can be added when word-specific memory fails or, in the case of novel forms like *spling*, does not exist.

A system where X is not just linked to Y, but all the features of X are linked to all the features of Y, may be called an associative memory, and connectionist pattern associators are a precise, systematic implementation of such a system. For example, Rumelhart and McClelland's pattern associator consists of an input layer representing the stem, an output layer representing the past tense form, and weighted connections between every input unit and every output unit. Each node corresponds to a sequence of phonological features, such as a high vowel between two stop consonants, or a back vowel followed by a nasal consonant at the end of a word. When a set of input nodes is activated, each node sends its activation level, multiplied by the link weight, to the output nodes it is connected to. Each output node sums its weighted inputs, compares the result to a threshold, and probabilistically turns on if the threshold is exceeded. The output form is the word most compatible with the set of activated output nodes. The model learns when a "teacher" presents it with input stems and their correct outputs. It adjusts the weights on the connections and the thresholds on the output nodes by small amounts in directions that reduce the difference between its own output and the correct output supplied by the teacher. The model learned irregular past tense forms, successfully generalized them to novel irregular stems, and learned irregulars in manners reminiscent of children, all because it is designed to capitalize on the redundant, hypersimilar patterns that irregular verbs display.

As a more realistic model of the memory in which irregular verbs are stored, associative networks like the Rumelhart-McClelland model can easily be incorporated into theories invoking minor rules. The verbs to which, say, the *i-a* rule applies would be stored in a pattern associator. This much is not controversial. Pinker & Prince (1988, 1992) make the somewhat more controversial step of proposing that such networks can model not only the list of verbs undergoing irregular alternations, but the alternations itself; no irregular rule such as Lowering Ablaut would be needed. The pattern associator would pair *sing* with *sang*, not *sing* with *i-a*; the *i-a* alternation is merely the strongest of a large ensemble of connections between stem components and past components.[48] But we will not discuss that step in this paper, because there is a more fundamental controversy on the table: what to do with *regular* past tense forms, which virtually all linguists attribute to a mental rule.

If a pattern-associator memory is already needed to link *sing* to *sang* and generalize

---

[48]Note that this proposal, while differing from that of Halle and Mohanan, is consistent with those of many other generative linguists. Jackendoff (1975), Aronoff (1976), Lieber (1980), Perlmutter (1988), and Spencer (1990) argue that semiregular patterns like those found among irregulars are not the product of real rules in the sense of mental operations, but of "redundancy rules," namely statements of common patterns reinforced by similar items in memory. Pinker and Prince (1992) note that redundancy rules do just what pattern associators do, and hence the latter are a good model of the former.

spling to *splang*, why not use it to link *walk* to *walked* and generalize from *wug* to *wugged*, eliminating the need for a separate rule that adds *-ed* to a symbol for the verb stem? This is exactly what Rumelhart and McClelland argued for: regular and irregular verbs are linked to their past tense forms in a single pattern associator, and differ only in degree, not in kind. Irregular verbs belong to classes of different sizes displaying different degrees of systematicity; regular verbs simply constitute a much larger, much more systematic class of the same kind. Regulars and the various subclasses of irregulars would simply carve up the space of English verb sounds among them. One region would be associated with *-ang*, another would be associated with *-ound*, and many large regions would be associated with *-ed*.

The appeal of this proposal is obvious. The memory mechanism uncontroversially needed to capture irregular patterns is parsimoniously pressed into service for regular verbs. Moreover, we would not have to try to explain how children learn a regular rule -- there is none. Nor would children need to classify verbs as regular or irregular. We will refer to this proposal as the Pattern Associator hypothesis. Note that while it happens to be embraced by most connectionists working on language, it is not an inherent part of connectionism. There could be separate connectionist networks for irregular and regular mappings, the latter representing all verbs by a single node rather than a set of nodes for phonological features. Thus the arguments discussed in this paper are not directed at connectionism itself, just at the Pattern Associator hypothesis professed by most connectionists.

Because of the apparent economy of the Pattern Associator hypothesis, it is important to examine the evidence that has led theorists to maintain that generalizations of the regular pattern are performed by a mental concatenation rule. We review this evidence in the next section, then consider the connectionists' reply before describing experiments that can decide between them.

Evidence That Regular Generalizations
are Computed by A Symbol-Concatenation Rule

The hallmark of a symbol is that it can uniformly represent an entire class of individuals, suppressing the distinctions among them. Moreover, the class can be open-ended; possible members of the class are represented identically to existing ones. In the case of the past tense rule, the fact that the rule concatenates the suffix *-ed* to "V," a symbol standing for the verb stem, means that the regular suffix can be productively attached to any verb, whether familiar or unfamiliar, similar or dissimilar to remembered regular verbs. The only exception is when memory already contains an irregular form listed for a given verb stem. In that case the system blocks application of the general rule to that stem. (Different versions of this principle have been proposed, variously called the Uniqueness, Blocking, and Elsewhere priniciples; see Marcus, et al., 1992 for a review.) Furthermore, if memory has a pattern-associator component, then we might expect any verb that is similar to stored irregulars, and not just irregulars themselves, to evoke irregular alternations and hence tend to suppress regular suffixation. But aside from a competition effect from irregular verbs and their phonological patterns, regular suffixation should be omnipotent, applying to any word bearing the "Verb" symbol. This offers the crucial empirical contrast to irregular alternations, which are linked to specific memorized verbs and their patterns.

The evidence for regular suffixation by rule, then (as summarized by Pinker and Prince, 1988, 1992; Pinker, 1991) consists in showing that in a large and diverse set of circumstances in which access to a memorized past tense form or pattern is prevented, held together only by the fact that a word bears the symbol "Verb," people readily supply the regular suffix. There are at least 20 such circumstances, falling into several classes corresponding to ways in which access to information in associative memory can fail: there may be no memory entry (or similar entry) at all; there may be competing entries; the grammatical mechanisms that allow information in memory entries to be passed to the word may be systematically disabled; the person's memory system may not be functioning properly. But despite this lack of memory access, generally revealed by a failure to apply the irregular form, the word is not left inflectionless: the regular rule, applying to the all-embracing, exemplar-independent symbol "Verb" (or "Noun"), succeeds in affixing the regular suffix in every case. (The arguments to be presented are independent of whether people in fact store common regular past tense forms in memory; whether they do or not, use of the regular affix does not *depend* on such stored forms.) The circumstances are listed in Table 1, where they are grouped for convenience into similar types.

Table 1:
Circumstances in Which Memory Patterns are not Accessed and Regular Inflection is Applied

Circumstance	Kind of Word	Example
Lack of Entry or Similar Entries in Memory:		
1. No root entry	Novel words	*snarfed, wugged*
2. Weak entry	Low-frequency words	*stinted, eked*
3. No similar entries	Unusual-sounding words	*ploamphed, krilged*
Competing Entries or Similar Entries in Memory:		
4. Competing root entry	Homophones	*lied/lay, hanged/hung*
5. Competing similar root entries	Rhymes	*blinked, glowed*
Entry is Not a Canonical Root:		
6. Rendering of sound	Onomatopoeia	*dinged, peeped*
7. Mention versus use	Quotations	*"man"s, "woman"s*
8. Opaque name	Surnames	*the Childs, the Manns*
9. Foreign language	Unassimilated borrowings	*latkes, negligées*
10. Distortion of root	Truncations	*synched, man.'s, OXes*
11. Artifical	Acronyms	*PACs, MANs*
Root Cannot Be Marked for Inflectional Feature:		
12. Derivation from Different Category	Denominal verbs	*high-sticked, spitted*
	Deadjectival verbs	*righted*
	Nominalized conjunctions	*ifs, ands, buts*
Features Cannot Percolate from Root to Whole Word (Exocentrism or Headlessness):		
13. Derivation via different category	Denominal nominalized verbs	*flied out, costed*
	Nominalized denominal verbs	*wolfs, gooses*
14. Derivation via name	Eponyms	*Mickey Mouses, Batmans*
	Products	*Renault Elfs, Top Shelfs*
	Teams	*Toronto Maple Leafs*
15. Referent different from root	Bahuvrihi compounds	*sabre-tooths, low-lifes*
	Pseudo-English	*walkmans*
16. Lexicalization of a phrase	Nominalized VPs	*bag-a-leafs, shear-a-sheeps*
Memory Failures:		
17. Children	Overregularizations	*comed, breaked*
18. Normal speech errors	Overregularizations	*comed, breaked*
19. Alzheimer's Disease	Overregularizations	*comed, breaked*
20. Williams Syndrome	Overregularizations	*comed, breaked*

Lack of Entry or Similar Entries in Memory

1. The verb has no past tense entry in memory. The fact that children (Berko, 1958) and adults supply regular past tense forms for verbs they have never before heard in the past tense (e.g., *snarf*, *mung*, *wug*) is the most familiar rationale for a rule.

2. The verb could have only a weak past tense entry in memory. Since memory strength depends on frequency of encounter, if inflection is dependent on memory retrieval, lower-frequency past tense forms should be less natural-sounding and harder to produce than high-frequency ones. Pinker & Prince (1988) pointed out that this is true for irregulars but not regulars: rare *bade*, *slunk*, and *stridden* sound stilted relative to their stems; rare *infarcted, eked*, and *stinted* do not. Quantitatively, low-frequency regular past tense forms are not rated worse (Ullman and Pinker, 1993; Pinker, 1991) nor produced more slowly (Prasada, Pinker, and Snyder, 1990; Seidenberg & Bruck, 1990; Daugherty & Seidenberg, 1992) than high-frequency ones.

3. The verb has no similar entries in memory. If memory has a pattern-associator component, lacking a past tense entry is not fatal; a past form may automatically be assembled from similar verbs with overlapping stem representations. Such analogizing is essential for generating novel irregular roots: people can inflect *spling* as *splung*, but only because of its similarity to *string, sling*, and so on; *nist* and *vin* do not yield *nust* and *vun* (Bybee & Moder, 1983; Prasada & Pinker, 1993). In contrast, people reliably supply, and give high ratings to, regular past forms for verbs that do not resemble any existing regular verb, such as *ploamph* and *krilg*.

Competing Entries or Similar Entries in Memory

4. The verb has a competing entry in memory. In associative memories, alternative mappings compete for specific patterns of input features: if a verb's sound is linked to one kind of past tense form, it will be linked to that form across all instances. But the regular form can overcome this competition, and apply to verbs whose sounds are reliably linked to an irregular form. There are several homophone sets in English where one verb takes an irregular form and the other takes a regular form; this would be impossible if both forms were generated as associations to the phonology of their stems, which are identical (Pinker & Prince, 1988). Examples include *lie-lay* (recline) and *lie-lied* (prevaricate); *hang-hung* (suspend) and *hang-hanged* (execute, in the standard dialect); *fit-fit* (intransitive) and *fit-fitted* (transitive), and *meet-met* and *mete-meted*.[49] These forms can coexist in rule theories because the two members are represented not as pure sounds but as abstract *entries* or *roots*, to which sounds are linked; the irregular past form is linked to one entry and not the other, and the regular rule is free to apply to the second.

5. The verb has competing similar entries in memory. Again, because of the associative nature of memory, families of similar listed entries should pull a new entry toward its associated pattern. But the omnipotent regular past tense process easily escapes this attraction: *blink* is inflected as *blinked*, not *blank* or *blunk*, despite competing *drink*, *stink*, *sink*, *slink*, and so on. Indeed, in every irregular territory in phonological space, there are interloping regulars: *jutted* next to *cut, shut*; *needed* next to *bled, bred*; *seeped* next to *slept, wept*; *glowed* next to *grew, blew*, and so on (Pinker & Prince, 1988).

Note that this does not imply that regular verbs are not subject to attraction from competing irregular patterns. Regulars similar to irregulars are produced more slowly (Daugherty & Seidenberg, 1992) and are subject to frequency effects (Ullman & Pinker, 1993; Pinker & Prince, 1992), in contrast with regulars that do not resemble irregulars. Indeed such effects are inevitable given that a verb cannot be recognized as being regular until after it has been checked against memorized irregulars and found to be absent; similarity to irregulars in an associative memory will yield a weak false-positive match signal. Our point is that the regular form generally prevails for verbs rhyming with

---

[49]In many regular/irregular homophone pairs, the two senses differ in transitivity or inherent aspect (the temporal contour of the event). For example *He fitted me with dentures* refers to the accomplishment of a change of state; *The dentures fit* refers to that state. *Hanged* is instantaneous; *hung* a state or achievement of a state. *Shone* can be used only with intransitive *shine* (*The sun shone/*Melvin shone his shoes*); *shined* can be used with either version. A possible explanation is that when a past tense form is stored in memory because of its idiosyncratic morphology, it can register idiosyncratic semantic information as well. Thus its meaning can contain not just generic pastness, but a representation the particular aftermath of that verb's event (e.g., whether its aftermath involves a change of state). When the verb is extended to have a related meaning with a different temporal contour, the stored semantics of the irregular past tense entry may not match that of the new event, which may in turn cause the irregular form to feel inappropriate. The speaker could then apply the regular suffix to the new form, because its semantics involves the composition of pure pastness with any event or state.

irregulars despite this pull. Not only are forms like *blinked* and *glowed* common in English, but in experiments, subjects usually prefer the regular form even when the attraction to irregular patterns is at its maximum, as in *spling* (Prasada and Pinker, 1993).

The Entry is Not a Canonical Root

Words are not represented in the mental dictionary as haphazard collections of information, but are generally based on entries of a standard format that can be called "canonical roots." This format consists of: an "address" or distinct identity as a word in the language, a part-of-speech category, relevant subcategory features (e.g., transitive or intransitive for verbs, count or mass for nouns), a semantic representation, and a phonological representation. The phonological representation must conform to a canonical template for possible phonological words in the language (McCarthy and Prince, in press) but are otherwise arbitrarily paired with the meaning for that word (de Saussure, 1960).

Complex words may be formed out of roots by processes of "derivation" or "word formation" (e.g., *unmicrowaveability*, *Nixon-hater*), but it is the roots that are the most basic units of lexical memory storage. And crucially, information about irregular forms (e.g., "my past tense form is *stuck*") consists of a link to a lexical root, not just to any kind of object that can function as a word. This fact gives rise to a number of other cases where information about inflectional form cannot be taken from memory.

6. The word lacks a canonical root in memory. To fill lexical gaps, people sometimes use sounds that are outside the system of canonical roots. One example is onomatopoeic words, which are conceived of not as standard-format arbitrary phonological objects but as renderings of the sounds in the world they refer to.[50] Onomatopoeic forms

---

[50]Of course onomatopoeic forms are usually conventional and consistent with the phonological constraints of the language, but these constraints may be violated (e.g., no word in English has the sequence found in *oink*), and crucially, speakers *perceive* onomatopoeia as representing language-external sounds, regardless of how tenuous and language-dependent the resemblance in fact is.

may be used as verbs, as in *My car tends to ping*, or as nouns, as in *I heard another bang*. Though they need past tense and plural forms, these words have no canonical roots, and hence cannot tap into the system of roots and their associated irregular forms that encourage irregular analogizing. Thus despite the fact that all the existing English verb roots ending in *-ing* are irregular (Pinker & Prince, 1988), which should make analogy especially strong, onomatopoeic verbs ending in *-ing* are regular, as are other seemingly analogizable onomatopoetic forms: *dinged/*dang/*dung, pinged/*pang/*pung, zinged/*zang/*zung, peeped/*pept, beeped/*bept.*

7. The word is mentioned rather than used. Another example of a sound that is used as a word but that lacks a canonical root is a quotation. A quotation bears no necessary resemblance to canonical words in the language, but directly reflects a stretch of sound that someone else has said or written. Of course, they may happen to be canonical words in the language, as in (1(c)), but this is perceived as coincidental.

(1)

a. He shouted "Dwat!" and then vowed to get that "wabbit."
b. Wayne and Garth's reaction was an emphatic "Schwing!"
c. Because of his fondness for "trivial" and "obvious" and "vacuous," his prose has an unnecessary polemical tone.

For the same reason that onomatopoeic forms fail to access the associative memory system for genuine roots, quoted nouns fail to do so. Information about stored plurals, in particular irregular plurals, is not tapped, but the regular applies:

(2)

While searching for examples of sexist writing, I found three "every man"s/*"men" on page 1.

8. The word is based on a name. In modern English, proper names are opaque stretches of sound. Of course, some of them happen to be homophonous with roots (e.g., *Baker, Green*), generally because historically they began as mnemonic descriptors of the person, but this is now conceived of as coincidental; a person could also bear the name

Dweezil or *Yazstrzemski* or *Bftsplk.* As in onomatopoeia and quotations, they are mentally represented as stretches of sound, not canonical roots, and hence do not evoke homophonous or similar canonical roots. Thus people regularize them:

(3)

We're having Julia Child and her husband over for dinner. You know, the Childs/*Children are really great cooks.

Sy and his brother Don, the Newhouses [unvoiced "s"] own a vast publishing empire.

Why has the second half of the 20th Century produced no Thomas Manns/*Menn?

9. The word is an unassimilated borrowing. Another example of words lacking roots consists of borrowings like *chutzpah*, *shwarma*, *negligée*, and *karaoke*, which are recognizably not canonical English words, but sounds taken from other languages. Despite their independence from a system of memorized roots, borrowed nouns easily receive regular inflection, as in *negligées* or *latkes* (though the case that memory is being bypassed would be even stronger if there were a competing irregular homophone or similar-sounding nouns). Among verbs, it is striking that virtually every French and Latin loan verb is regular (Pinker & Prince, 1988). Presumably this is because Latinate verbs are typically polysyllabic, with stress on the second syllable, and the canonical template in English is a monosyllable or a disyllable with initial stress (McCarthy & Prince, in press; Pinker, 1989; Pinker & Prince, 1988). Noncanonical borrowed verbs, even if similar to existing clusters of irregulars, are all regular: *derided/*derode* (cf. *ride/rode, stride/strode*; *succumbed/*succame* (cf. *came*).

Of course, many borrowings may get assimilated into canonical words in the language, that is, speakers assign them canonical lexical entries. For example, in English, many French loans like *promise* have taken the canonical English stress pattern and are treated in all ways identically to the native vocabulary (Pinker, 1989; McCarthy and Prince, 1990, in press.) The few Latinate borrowings that have been given irregular pasts like *cost*, *quit*, and *catch*, are clearly assimilable to the English template. Several factors determine whether speakers convert a loan word to a canonical word in their language or represent it as a special sound. These probably include the degree to which the loan word can easily be distorted to fit the canonical phonological template of the language (McCarthy and Prince, 1990), and metalinguistic factors such as a speaker's awareness of the actual loan status, their sensitivity to the phonological templates of the borrowed-from language, and perhaps even the degree to which the referent of the loan word is a typically namable entity in the borrowing culture. Regardless of what these factors are, there should be a representational difference between unassimilated borrowings, which should resemble onomatopoeia, quotations, and names in being sounds that do not have a root or canonical lexical entry, and assimilated borrowings, which should be no different from any other root in the language. (This distinction will be important in interpreting Experiment 2.)

10. The word is formed by distorting a root. Words perceived not to be roots, but truncations of roots (often done for whimsical reasons), should not tap into memory for roots; the regular applies. Thus we get *synch'd* (from *synchronize*), not *sanch*. Similarly, if in some specialized argot or jargon a *manufacturer* was referred to as a *man.*, it seems likely that the plural would be formed as *mans*, not *men*; if *Oxygen* were referred to by the truncation *Ox* and different kinds then referred to, the plural would surely be *Oxes*, not *Oxen*.

11. The word is formed by artifical means. Acronyms are not roots, but are formed by an artificial method from the orthographic rendering of roots. Nonetheless they undergo regular suffixation, as in *PACs* (Political Action Committees). Indeed even if the acronym matches a stored irregular past form, the form is unusable, and regular suffixation would apply. If there were a class of substance called a Moderately Active Narcotic with abbreviation *MAN*, its plural would obviously be MANs (and not MEN -- indeed presumably an acronym OX would not be pluralized irregularly as OXen even though that would preserve the identity of the acronym).

The Root Cannot Be Marked For the Inflectional Feature

12. The word is derived from a root of a different grammatical category, which is incapable of bearing the inflectional feature in its entry. A denominal verb is a verb that is sensed by speakers to be derived from or based on a noun. Denominal verbs uniformly have regular past tense forms, even when homophonous with an irregular verb. Examples are shown in (4). See Pinker and Prince (1988), Kiparsky (1982), Kim, Pinker, Prince, & Prasada (1990), and Kim, Marcus, Pinker, Hollander, and Coppola (1992), for more examples, documentation of some of their printed sources, and experimental data showing that college students, non-college-educated adults, and 4-8-year old children reliably regularize denominal verbs.

(4)

He *spitted/*spat* the pig. (put on a spit)
He *ringed/*rang* the city with artillery. (formed a ring around)
Martina *2-setted/*2-set* Chris. (beat in two sets)
He *high-sticked/*high-stuck* the goalie. (hit with a high stick)
He *braked/*broke* the car suddenly. (applied the brakes)
He *steeled/*stole* himself for the ordeal. (made like steel)
He *meaned/*meant* both lists of numbers. (calculated the mean of)
He *stringed/*strung* the sugar snap peas (removed the string)

A noun root like *stick* cannot have an irregular past tense associated with it because a noun cannot have any past tense associated with it, the notion of "past tense" making no sense for a noun. The regular inflectional rule, being the default, is fully available -- indeed, is the only way -- to inflect such derived verbs.[51]

For similar reasons, a verb derived from an adjective should take the regular suffix (even if homophonous with an irregular verb): the past tense of *right the boat*, i.e., "to set

---

[51]The difference between homophones and denominal verbs is that for homophones, a regular form for one of the homophones is possible, but whether it in fact exists is an accident; the homophones could correspond to two entries that are both irregular, as in *rang the bell* versus *wrung out the shirt*. For denominal forms, one of the two roots, namely the one that is not a verb, *must* take the regular form.

right," is indeed *righted*, not *rote*. Likewise, a noun derived from other grammatical categories should easily take the regular plural suffix: *Their typical reaction is a blizzard of ifs, ands, or buts*. (In this case there is no competing English irregular sound pattern, but as we shall see, there can be in other languages.)

The regularization of irregular-sounding words lacking appropriate roots is part of a more general phenomenon. People's grammatical systems are built with a specific mechanism, "percolation," by which the information stored with a root can be passed on to any complex word containing that root (Williams, 1981; Selkirk, 1982; Kim, et al., 1991; Pinker & Prince, 1992.) In the constituent structure reflecting a word's derivation from more basic morphemes, one of these morphemes is generally the *head* of the word, and its properties percolate up to characterize the word as a whole. In English and many other languages, the head of the phrase is the rightmost element, so that when the constituent structure of a complex word is represented as a tree, the heads will be seen as chain of category labels on the right periphery.

(5)

```
 X
/ \
   X
  / \
     X
    / \
       X
```

Thus the head of *overeat* is the verb *eat*, so overeating is a kind of eating, and it is a verb (as represented by the topmost node), the same way *eat* is a verb. Similarly, a *workman* is a singular noun referring to a kind of man, not a kind of work.

(6)

```
       V                      N
      / \                    / \
     /   \                  V   N
  prefix  V                 |   |
    |     |               work  man
   over  eat
```

Crucially, the percolation conduit from the head to the top node of a word structure applies to *all* the information stored with the head. Not only does the grammatical category and meaning of the head percolate up (and features like gender, humanness, animacy, and inherent aspect), but for irregular words, the irregular past tense or plural form stored with the head percolates up as well. Thus the past tense of *overeat* is *overate*, and the plural of *workman* is *workmen*. This is also why we get irregularity preserved in novel forms like *out-sang, overshot, re-did, sawteeth, repairmen*, and *superwives*.

But some words are headless, or "exocentric": they differ in some property from their rightmost element, requiring that the usual pipeline of information percolation from head to top node be blocked. Denominal verbs, for example, are necessarily headless. Consider the structure of the verb *ring* from *ringing the city*.

(7)

```
  V
  |
  N
  |
ring
```

Since the whole word, represented by its topmost label, is a verb, but the element it is made out of, *ring*, is a noun, it must be headless -- if the noun *a ring* were its head, *to ring* would have to be a noun, too, which it is not. Though we have just noted that the noun *ring* cannot be listed with a past tense form for semantic reasons, even if it could be, the denominal verb *to ring* could not inherit it, for without a head and its associated data pipeline, the stored form *rang* could not bubble up to attach to the whole word. Thus even if a word has a root that *does* bear the relevant features in memory, if the word is headless, the features remain confined with the root, and are unavailable for application to the whole word.[52] This gives rise to several other regularization circumstances.

Features Cannot Percolate from Root to Whole Word (Exocentrism or Headlessness)

13. The inflectional information stored with an irregular root in memory is unavailable to the word as a whole because derivation via a different category makes the word headless. Consider the baseball term *to fly out*, meaning "to hit a fly." The noun it comes from, *a fly* (sometimes *fly ball* or *pop fly*) is itself clearly based on the simple verb root *to fly* ("to proceed through the air"). Therefore *to fly out* does have a verb root, attached to the irregular past *flew*. However, the step in the derivation that derives the verb *(to fly out)* from the noun *(fly ball)* yields an exocentric structure, as does the step in the derivation that derives the noun *(fly ball)* from the root verb *(fly)*.

(9)

```
 V
 |
 N
 |
 V
 |
fly
```

Therefore, access to the memorized *fly-flew* connection is blocked, and the most common past tense form is *flied*. Similar examples include *the doctor casted his leg* (= "put

---

[52]Another way of capturing this phenomenon is to posit that no verb is truly headless. When a verb is derived from a noun, it is headed by a phonologically empty derivational suffix, as in the following:

(8)

```
    V
   / \
  N   V
  |   |
ring  Ø
```

The noun would still not be the head of the verb; the silent suffix would be. The suffix would be categorized as a verb and hence would determine the category of the whole word in a fashion parallel to that of other derived words. See Wunderlich (1986), Olsen (1990), Neeleman and Schipper (1992), Brinkmann (1993), and Pesetsky (in press), for arguments in favor of this treatment.

a cast on," ultimately from the verb *to cast* but only via the noun *a cast*), *Our grant administrator costed out the equipment requests for us* (= "ascertain the cost of," from verb *to cost* but only via the noun *the cost*), *Alan grandstanded to the jury* (= "play to the grandstand," from the verb *stand* but via the noun *a grandstand* (see Kiparsky, 1982).[53]

For similar reasons, nouns derived from verbs are exocentric, and take regular plural forms even if homophonous with an irregular noun (indeed, even if ultimately derived from that noun). For example, the denominal verb *to wolf* means "to gobble like a wolf." Like many verbs, it can be nominalized, or turned into a noun, referring to an instance of wolfing.

(10)

```
  N
  |
  V
  |
  N
  |
wolf
```

(11)

With a couple of quick wolfs/*wolves, Arnold consumed his entire lunch.
While we're at the conference in Maine I hope we'll have time between sessions for a couple of quick fishes/*fish.
During the last faculty meeting Harold gave Sam a few knifes/*knives in the back.
He reached over and gave her a couple of quick playful gooses/*geese.
I need to find some examples of perfume ads; can you take this magazine and do a couple of quick leafs/*leaves through it?

14. Derivation via a name blocks the percolation of information from the root entry. As mentioned above, pluralized surnames like *the Childs* regularize because a surname is not a canonical root. But the phenomenon is even more general: names that do have

[53]The complex structure of these seemingly simple verbs makes them liable to misanalysis by some speakers, who may short-circuit the derivation and assume that the derived verb is directly based on the root. Such speakers are predicted to retain the irregular form, explaining apparent counterexamples to the regularization phenomenon, such as the occasionally-heard *flew out*. See Kim, et al., (1991) for empirical documentation of this effect.

canonical roots regularize, too. For example, the proper name *Mickey Mouse* is obviously based on the noun root *mouse* with its irregular plural *mice*. But when the name is converted into the eponymous noun *a Mickey Mouse* ("simpleton"), the irregular is unavailable and the regular steps in: *I'm sick of dealing with all the Mickey Mouses/*Mickey Mice in Administration.* Eponyms can also come from dolls, theatrical roles, and titles; in addition, the brand names of products and the nicknames of sports franchises can be derived from irregular noun roots and then converted back into nouns. In all cases they tend to regularize (See Kim, et al., 1992, for documentation of the effect in children.)[54]

(12)

I just bought Joshua two more Batmans/*?Batmen for his collection.
He's the best of the Mac the Knifes/*Knives in this run of Beggar's Opera.
This is the best of all the Pretty Womans/*Women that have been recorded.
Hollywood has increasingly been relying on sequels, as
the three Supermans/*Supermen and two Batmans/*Batmen attest.
Renault Elfs/*Elves (cars)
Spectrums/*Spectra ("Spectrum" brand bicycles)
Top Shelfs/*Shelves ("Top Shelf" brand frozen dinners)
The Toronto Maple Leafs/*Leaves (hockey team)[55]

The effect is a direct consequence of the fact that names are not common nouns, and therefore must be represented as a different kind of lexical category from N. As we saw earlier, names need not, in general, be canonical roots. Say a stretch of sound that functions as a word but lacks a canonical root (names, borrowings, onomatopoeia, quotations, acronyms, truncations) bears a nonspecific symbol "X" as its category label in the mental

---

[54] Occasionally product names are given irregular forms like *Vaxen* (Vax computers) and *Macinteesh*, but these are widely recognized as self-conscious linguistic humor, a oft-noted characteristic of the "hacker" culture from which these examples originated (Raymond, 1992).

[55] The *Minnesota Timberwolves* go the other way. Presumably this is because team nicknames can be taken directly from plurals when semantically appropriate (e.g., the entire team is being identified with a pack of wolves), inheriting any irregularity, just as nicknames can come from mass nouns like the *Miami Heat* or *Orlando Magic*. In contrast, when a name is clearly based on the singular, it does regularize: In November 1992, Boston Globe columnist Alex Beam referred to several navy planes as *Sea Wolfs*. Similarly, for the Leafs, the team is not being identified with foliage, but each player is identified with Canada's national symbol, the maple leaf.

dictionary. When it is converted to a noun referring to the generic object that is pluralized, the noun is exocentric:

(13)

```
      X                N
      |                |
 Julia Child           X
                       |
                  Child (s)
```

When the name is itself transparently based on a noun (say, *elf*), the structure is exocentric for two reasons (it does not have the same category symbol, N, all the way up), and hence percolation-proof; links to irregular forms like *elves* are trapped in the entry and the regular applies.[56]

(14)

```
  N        X        N
  |        |        |
 elf       N        X
           |        |
          Elf       N
                    |
                  Elf (s)
```

15. The semantic interpretation of a compound precludes using the features of the root. The headlessness account also explains seemingly unrelated cases in the plural system where there is no category change. Consider *low-life* -- not a kind of life, but a kind of person, namely one who leads a low life. This is an example of a "bahuvrihi" compound, a compound that refers to an object characterized by *having* rather than *being* the referent of its rightmost morpheme. Recall that the head-to-top-node pipeline cannot be blocked for just one kind of information; if it is blocked for one thing, nothing passes through

---

[56] On semantic grounds, too, the structures must be exocentric, for the eponym *a Mickey Mouse* is not a kind of mouse in the same way that *a workman* is a kind of *man*. Moreover, in the case of eponyms, it could be argued that the derivation is exocentric for three reasons, because it the name is made into a NP before being converted back into a noun: in English a proper name behaves syntactically like an entire noun phrase, not like a common noun (**the Donald; *a Gary; *I'm going to speak to happy Fred*).

automatically. That means that there is no way for the irregularity of a headless word to percolate up: If *low-life* cannot get its referent from *life*, it cannot get its plural from *life* either. Though memory for stored past tense forms, in this case an irregular one, is unavailable, the regular rule steps in by default; hence *low-lifes*, not **low-lives*. Other examples of regularized exocentric compounds (sometimes co-existing with the irregular form) include *still lifes*, *bigfoots* (monsters), *flatfoots* (policemen), *tenderfoots* (wolf cubs or cub scouts), *goofyfoots* (left-footed surfers), *hotfoots* (practical jokes), *bigmouths* and *loudmouths* (blowhards), *saber-tooths* (cats), and *sweet tooths* (taste for sugar).

A related phenomenon occurs in the now-generic product name *walkman*, whose plural, according to *Newsweek* (August 7, 1989, p. 68), is commonly *walkmans*. The word is not a bahuvrihi compound or any other standard kind of English compound; it was coined by the Sony corporation, an example of the pseudo-English common in Japanese products and advertising. Its structure, obscure to an English speaker, is hard to analyze as referring to a kind of man; hence it would be treated as exocentric, and thus regularized for the same reasons that bahuvrihi compounds are. See Kim, et al. (1992) for data showing that children tend to regularize exocentric compounds.

16. The word is formed from a phrase, in which a root is inaccessible. Occasionally people coin nouns from verb phrases to refer to tools that have the function specified by the VP. Again the structure is headless, hence regularized.

(15)

```
      N
      |
      VP
     / \
   V     NP
   |    / \
 bag   a  leaf
```

(16)

bag-a-leafs
protect-a-childs (car seats)
repel-a-deers
shear-a-sheeps (razors)

Memory Failures

In any case in which a person's memory retrieval mechanism fails to deliver a stored irregular form but that person's grammatical system is intact, a regularized version of that form should be available. There are four documented cases.

17. Childhood overregularizations. Children have not heard past tense forms as often as adults, and so are expected to have less reliably-accessed memory entries for them, especially the lower-frequency ones. Nonetheless they succeed in marking them for tense, in overregularization errors like *holded* (Marcus, et al., 1992).

18. Speech errors. Normal adult memory is not perfect, either, and indeed, overregularizations occur as adult speech errors (Stemberger, unpublished data; cited in Marcus, et al., 1992).

19. Alzheimer's Disease. Alzheimer's patients characteristically have word-finding problems, a part of their general memory impairment; but they perform virtually perfectly on *wug*-tests, and often overregularize irregular verbs (Ullman, Hickok, & Pinker, 1992).

20. Williams Syndrome. This is an unusual form of retardation involving hyperlinguistic abilities, including excellent grammar, but deviant word retrieval, involving a reduction of the normal tendency to favor high-frequency words. The syndrome can be accompanied by high rates of overregularization errors (Bellugi, et al., 1989; Pinker, 1991; Marcus & Kelley, unpublished data).

Though there are many different circumstances in which information from stored roots is suppressed but regular affixation applies, it is important to note that the regular is *not* merely used whenever a form is novel, unusual, extended, or abstract. Lexical information about inflectional form can be passed to novel words in a number of circumstances, summarized in Table 2: (1) when a new word is perceived as a lexical root, and is similar enough to other roots that their associated inflected forms can transfer to it (e.g., *spling-splung*); (2) when a new word is formed by prefixation of an irregular verb, leaving the verb in head position, as in *overate, out-sang, overshot, re-did, re-broke, overshrank, underflew*, and *unhung*; (3) when a new compound word is formed with an existing irregular in the rightmost, head position, as in *workmen, repairmen, coalmen, superwomen, muskoxen*, and *milkteeth*; (4) when a word is used metaphorically or in some extended sense, as in *men* (chess pieces), *leaves* (of a book), and *Freud's intellectual children*, *oilmice* (Chinese peasants who scavenge oil from unguarded wells), *sawteeth, snowmen, metrical feet, three feet long*, and so on; (5) when a word is used in an idiom or collocation, particularly common with "light verbs" such as *cut a deal, took a leak, caught a cold, blew him off, put him down, came off well, went nuts*, and so on.

Table 2:
Novel Circumstances in Which *Irregular* Inflection Applies

<u>Root in Head Position:</u>		
1. Simple root	Irregular-sounding words	*splang, sprunk*
2. Head of derived word	Prefixed verbs	*out-sang, re-broke*
3. Head of compound	Endocentric compounds	*sawteeth, workmen*
4. Semantic extension	Metaphors	*(chess) men, (intellectual) children*
5. Idiom	Light verbs	*blew him off, took a leak*
<u>Root in Nonhead Position:</u>		
6. Combinations of stems	Lexical compounds	*mice-infested, teethmarks*

There is an especially interesting circumstance in which irregular forms may be used, because of their status as memorized roots, but regular inflected forms may not. Kiparsky (1982) has noted that irregular plurals easily appear inside compounds as the left-hand (non-head) member, as in *mice-infested, men-bashing, teethmarks*, and *oxen-yoke*) whereas otherwise similar regular plurals do not (**rats-infested, *guys-bashing, *clawsmarks, *horses-yoke*). This effect has been documented experimentally for 3-5 year old children by Gordon (1985) and for adults by Senghas, Kim, Pinker, and Collins (1991). The explanation is that this lexical compounding process is a derivational rule that takes two stems from the mental dictionary and joins them together. Irregular forms are stems in the mental dictionary, and can be used in compounds, just like any other word. But regular forms are not stems in the mental dictionary, but complex products of a rule, formed outside of the lexicon, too late in the chain of processes for inclusion in the compounding operation.[57]

In fact, the respective circumstances in which regular and irregular inflection can apply are not predicted by some global function of similarity, but are predicted in very precise ways by the theory of rule application -- all from a very simple set of postulates, required independently of the facts about regular/irregular interactions reviewed above. The theory only states that (a) lexical information is stored in the form of canonical roots in the mental lexicon; (b) memory is associative, so that similar roots reinforce each other by overlap and foster analogies to similar new roots; (c) information stored with a root can apply to a word containing that root only if the root is the head of the word; (d) the mental grammar contains symbol concatenation operations which can apply to any instance of a syntactic category unless there is a competing form passed on from lexical memory. All the facts about regular and irregular generalizations (20 circumstances in which the regular form is generalized, and 5 where irregular forms are generalized) fall out of this theory, on the assumption that irregular forms are lexically-stored information and regular forms are rule products. In all and only the circumstances in which a word is sensed not to be headed by a stored irregular root (or irregular-sounding root) -- that is, by default -- the word is given a regular inflected form.

---

[57]By "stem" we mean a canonical uninflected word that may have been formed from a root by processes of derivational word formation (see, Selkirk, 1982; Aronoff & Sridhar, 1983). Thus *dancer*, which can appear inside compounds like *dancer-lover*, is a stem formed from the root *dance*.

Crucially, the symbolic rule theory gives a single explanation to the hegemony of the regular form in circumstances that would otherwise seem to have nothing in common: nonexistent roots (*wug*), unusual roots (*ploamph*), headless words of a variety of kinds (*flied out*, *low-lifes*, *Mickey Mouses*), rootless words (*pinged*, *synched*), without making the incorrect prediction that any form of novelty whatsoever triggers regularization. What these diverse words have in common is being instances of the same symbol. There need be no special mechanism that ensures that an unusual-sounding root like *ploamph* gets its suffix, with separate ad hoc mechanisms ensuring that common-sounding nonroots like *ping*, common-sounding headless roots like *fly out*, and all the rest get that suffix, too. Rather, they all earn *-ed* simply by virtue of all bearing the symbol V. We emphasize this to underscore the fact that people's use of regular inflection in all of these circumstances falls automatically out of the nature of the underlying process, rather than being a collection of unrelated contingencies acquired one at a time from hearing the relevant uses in speech and writing (which, in any case, would beg the question of why previous speakers of the language applied a single suffix to all of those cases to begin with). This huge gain in parsimony is the reason that invoking a symbol-concatenation operation, of the kind needed for syntax, is mandated for regular inflection.

Attempts to get Pattern-Associators
to Mimic the Effects of Rules

The challenge for associationist theories of language is to account for these diverse circumstances in which regular inflection can be applied productively. The Rumelhart-McClelland model was designed to display two of the phenomena: regular inflection of novel words like *wug* (Table 1: #1), and children's overregularizations like *breaked* (Table 1: #17). Pinker and Prince (1988) and Marcus et al. (1992) have shown that its explanation for children's overregularization is incorrect (discussed below). Presumably the model can regularly inflect low-frequency words, a third circumstance (Table 1: #2), since it can regularly inflect novel, that, is, zero-frequency words; whether it can do so with equal facility to high-frequency words is unknown.

The model does quite poorly, however, when attempting to inflect novel unusual-sounding verbs like *ploamph* (Table 1: #3) (Pinker & Prince, 1988; Prasada and Pinker, 1993). It does not reliably apply regular inflection to novel forms resembling irregulars (Table 1: #5) (Prasada and Pinker, 1993), but since this difference from humans is only quantitative, its ultimate significance is uncertain.[58] Effects of the analogue of memory failure in neurological impairments (Table 1: #19) are uncertain, but Marchman (in press) shows that a model similar to the Rumelhart-McClelland that is "lesioned" by eliminating connections shows selective decrements in its ability to apply *regular* inflection, exactly the opposite of the pattern found in Alzheimer's Disease patients. And because the model's input representation consists entirely of phonological information, it is incapable in principle of regularizing words whose phonological representations are the same as those of familiar irregular words: homophones of irregulars (Table 1: #4), derived words with roots of different categories (Table 1: #12), headless words (Table 1: #13-16), and rootless words (Table 1: #6-11) (Pinker & Prince, 1988; Kim, et al., 1991).

There have been many suggestions by connectionist theorists about how to deal with these problems. The most common suggestion attributes these problems to the model's input representation, "Wickelfeatures," and its two-layer architecture: input and output are connected directly, rather than being mediated by a hidden layer of nodes (e.g., McClelland,

[58]The model can inflect some *existing* regular words rhyming with irregulars (e.g., *needed*). This ability, however, may have come from a quirk of that particular implementation. Rumelhart and McClelland represented each word by activating not only its actual phonological feature nodes, but an additional set of nodes that represented distortions of the actual word. They referred to this scheme as "blurring," but curiously, each word received the same set of blurred features on every presentation, so it was not like random noise. But many of the nodes in the model do not correspond to any actual phonological sequence in English, so consistent "blurring" is a way of recruiting these otherwise useless nodes as surrogates for a verb's idiosyncratic lexical entry. Similar verbs -- indeed, even homophones -- with different past tense forms could then be distinguished by the connections to these nodes.

1988; Plunkett and Marchman, 1991; MacWhinney and Leinbach, 1991; Seidenberg, 1992). In fact, the evidence suggests that these factors are irrelevant. Egedi and Sproat (reported in Sproat, 1992) have performed the only focused comparison of two-layer and three-layer networks for inflection, comparing the Rumelhart-McClelland model to their own improved version, which had a hidden layer and an improved input representation and output process, and which they trained on the same set of items. Their model behaved virtually identically to Rumelhart and McClelland's: it failed to provide regular inflected forms for novel unusual-sounding words, and failed to overregularize except with the same unrealistic input regime that the original model needed. For the other circumstances in which rules provide regular forms for irregular-sounding words (homophones, headless and rootless words, etc.) there is no issue: if a model fails to represent nonphonological information like lexical identity, derivational structure, and roothood, it is incapable in principle of discriminating homophones displaying such differences, regardless of the number or wiring of hidden layers, just as an organism with a single kind of photoreceptor is incapable of discriminating isoluminant hues, no matter what brain circuitry we attribute to it.

A second technological response is to add a separate hardware feature to cope with each of the regularization failures. This was the approach taken in developing the MacWhinney and Leinbach (1991) model. First, irregular phonological patterns (generally rhymes) support analogization, but nonetheless regular affixation can often apply to individual words displying such patterns, (Table 1: #5, e.g., *blinked*). Thus MacWhinney and Leinbach built in two redundant phonological representations for words in the input, one for the rhyme, allowing the analogies, one for the whole word, allowing the regular exceptions. Second, irregulars are largely idiosyncratic in membership and form, but regular inflection predictably adds a suffix to a preserved stem regardless of its composition (Table 1, #3). Thus they added a second pathway to the model, tailor-made for regular suffixation, which connected every input directly to its counterpart in the output, supplementing the pattern-associator network. Third, some homophonous verb pairs can have different past tense forms (Table 1, #4). Thus they added nodes for the semantic features that are necessary to distinguish such homophones, such as features specific to the meaning of *ring* and specific to the meaning of *wring*.

But such efforts speak primarily to the engineering question, "what does one have to do to a connectionist network to get it to reproduce the facts of inflection," rather than the scientific question "what are the mental mechanisms that make inflection the way it is?" MacWhinney and Leinbach did not motivate their two lexical representations, or their two pathways, except as ways to get the model to work and as fixes to Pinker & Prince's (1988) individual criticisms. MacWhinney and Leinbach's solution to the regularization problems is particularly unrealistic as psychology: They examined the handful of pairs of English words that by historical accident are homophonous with different past forms (like *wring* and *ring*), and built-in innate features for just such pairs (not for all words). Moreover, the suggestion that the semantic features used as cues to discriminate homophones (Table 1: #4) can also be used to regularize headless and rootless words like *fly out* (Table 1: #6-16) is empirically inaccurate: it is not particular semantic features, or even degree of overall semantic unusualness, that makes an irregular verb regularize (see Table 2, and data from Kim, et al., 1991); it is the more abstract relational notion of headlessness or rootlessness. Finally, the fact that people immediately regularize headless and rootless words even in exotic circumstances like inflected quotations and truncations shows that the phenomenon has to be an automatic consequence of how regularization is computed, not a set of specific contingencies trained by exposure to actual regularized words from each circumstance. (See Kim, Marcus, et al., 1993, for evidence that children regularize denominal homophones of irregular verbs, even though they probably have not heard any in parental speech.)

Hare and Elman (1992) describe a standard three-layer model trained on 32 Old English verbs from each of 6 different classes of inflection. The input pairs each training verb with a code for which inflectional class the verb is from. Unlike earlier models, it could generalize novel words that do not resemble any words in the input to their Class Six, the default class. (Regularization of homophones, headless, and rootless words was not addressed). But this model, too, attains technical success without regard to psychological plausibility. Whereas the standard connectionists models of inflection are designed to compute the phonological form of the output in response to training exemplars consisting of a stem and its past form, this model is wired to activate one of six local output nodes corresponding to one of the six inflectional classes, in response to training exemplars consisting of a stem and its inflectional class. But children hear their parents produce past tense forms; they do not hear their parents label verbs with their inflectional classes. Thus one of the orginal advantages of the Rumelhart-McClelland model -- that it need not explictly be told which items are regular and which are irregular -- is lost. Furthermore, if the model is taken seriously as a psychological theory, it is tantamount to the strange claim that the human brain is innately hard-wired for the six inflectional classes of Old English.

Putting aside these fixes, are there any general principles by which connectionist models might duplicate the wide-ranging productivity of regular inflection? Only one has been proposed: type frequency, or the number (and thus variety) of words with regular inflection. ("Type frequency" refers to the number of different words in a class, each counted once; "token frequency" refers to the number of occurrences of a particular word.) In addition to the various properties distinguishing regular and irregular suffixation listed in Tables 1 and 2, there is one more: the regular form applies to the majority of nouns and verbs in English, irregular forms to the minority. This of course is the original reason that the *-ed* suffix is called "regular" for English, in the purely descriptive or statistical sense of "regular." Since connectionist pattern associators are specifically designed to pick up on the input-output statistics, this is the sense, and the only sense, in which they acquire *-ed* suffixation as a "regular" operation. The extent to which pattern associators generalize the regular suffix depends on its frequency in the vocabulary sample it is trained on.

Connectionist modelers have been explicit about this point. Rumelhart and McClelland (1986: p. 230-231) note,

> Because of the predominance of the regular form in the input, the network learns the regular pattern ...

They modeled the onset of overregularization in children by exploiting this fact: the model overregularized when, after first being trained on a small set of verbs, 20% of them regular, it was suddenly trained on a much larger set of words, 80% of them regular. Bybee (1991), in proposing a kind of connectionist model for inflection, claims

> All types of morphological patterns can be acquired by the same process -- the storage of items, the creation of connections among them, and the formation of patterns that range over sets of connections. The differences among [inflectional classes] are due largely to the number of distinct lexical items involved -- a big class is more productive and forms a stronger schema than a small class. A large class has a high type frequency, that is, the number of different words containing the suffix or pattern is high (p. 86-87).

Plunkett and Marchman (1991) demonstrated empirically that varying the ratio of regular and irregular items radically changes the performance of their network model. For example, they found that increasing the number of irregular types has the consequence that

> performance on the regulars deteriorates. This result can be partially, but not totally, predicted from changes in relative class size (p. 67).

Like Rumelhart and McClelland, they had to model overregularization by changing their model's diet from one lean in regulars to one rich in regulars (Plunkett and Marchman, 1990). Similarly, in the Daugherty & Seidenberg (1992) model, reducing the number of regular types substantially weakened the model's ability to generalize the regular pattern to novel words. But in a training regime in which regular verbs were a large majority, the model duplicated people's ability to produce regular inflected forms for low token frequency verbs needing no more time than for verbs with high token frequencies (Prasada, Pinker, and Snyder, 1990; Seidenberg & Bruck, 1990). Seidenberg (1992, ms p. 14) speculates,

> One of the important tests of future models of the past tense will be to determine if they generalize in appropriate ways. I myself doubt whether this will be a serious problem, assuming the model is trained in a way that faithfully reflects facts about the distribution of regular and irregular past tenses in the language. The system is overwhelmingly regular, and the weights will come to reflect this fact, making it likely that the regular past tense will be attached to almost any novel input.

Unlike the technological revisions, these suggestions are motivated by a real fact about English, namely that regular verbs and nouns are in the majority, together with genuine principles governing how connectionist models work. High type frequency leads to generalization in connectionist models for several reasons. Generalization occurs when a new item is represented on some of the same nodes that belong to previously encountered words that have been paired with a particular output during training. Thus when many words have been paired with a particular output, the connections between their overlapping nodes and the output are strengthened many times, so a new word represented on those nodes has a higher probability of evoking that output. Furthermore, the more words that have been paired with a single output, the more different nodes, on average, will be linked with strong weights to that output (depending on how diverse the words in the training set are), so the more likely it will be that some new item will overlap with some previous one and exploit its connections. Finally, the thresholds of the nodes representing the output will be lowered when it is frequently trained, making such responses more likely across the board. For these reasons, a suitably-configured and suitably-trained pattern associator might in principle be able to inflect unusual-sounding words like *ploamph*, owing to their diffuse similarity to the large and widely spaced set of regular training exemplars.

But as can be seen from the speculative nature of Seidenberg's comments, at present there is no evidence that high type frequency really is sufficient to allow connectionist pattern associators to generalize appropriately. In practice, there are several factors that can militate against it. Since in a pattern associator the same set of nodes has to represent the inflectional class of the verb (i.e., irregular versus regular) and the actual phonological composition of the verb from which the stem portion of the past tense form is recovered, the representations of the input have to be quite fine-grained. This defines a high-dimension input space, where new verbs' representations can fall into nooks and crannies that do not overlap previously-trained verbs and hence cannot co-opt their connections (Prasada & Pinker, 1993). Moreover, the presence of irregulars in the training set -- generally with very high token frequencies (Rumelhart & McClelland, 1986; Bybee, 1985; Pinker and Prince, 1988) -- means that many nodes will have to be inhibited from producing the regular output, competing with the regular mapping. That is, many of the nodes representing a new form like *ploamph* will inhibit the regular output because they have been trained to do so for *go*, *blow*, *grow*, and so on.

By itself, type frequency does literally nothing to help the models' problems with regularization of irregular sounding homophones, headless and rootless words, and so on, because the necessary morphological and lexical distinctions are not even represented and so the model is color-blind to them. If nodes representing semantic features, or even individual lexical distinctness, are added to the input representation as a way of approximating such information, generalization to unusual sounding words might be further curbed, as the model could learn to associate the regular ending with frequent words nodes or semantic nodes, absorbing connection strengths that would otherwise go to general patterns of phonological nodes.

Still, the important role of type frequency might allow one to give the benefit of the doubt to future connectionist models. Conceivably there might be an input representation and network configuration that, when processing a realistic input mixture (with thousands of regular words, most of low token frequency, and a few irregulars, most of high token frequency), could generalize freely to novel, unusual-sounding, low-frequency, and most irregular-sounding words. Furthermore, if a modeler employed some input representation representing a word's root status and morphological structure, the model might be able to exploit hidden nodes that represent all of non-irregular phonological space (if they developed in learning unusual-sounding words), and develop other hidden nodes that inhibited the all the irregular regions, and use them together to select the regular suffix as the default across all the phonological and morphological circumstances that call for it, without having to be trained on each one.

Is there any way of testing the type frequency hypothesis directly, rather than speculating about what future models might or might not do? Pinker and Prince (1988) and Marcus, et al. (1992) did one such test: they found that the onset and degree of children's overrergularization errors were not predicted by increases in the number, or the proportion, of regular verbs in the parental input or the child's own vocabulary, contrary to the predictions of Rumelhart and McClelland (1986) and Plunkett and Marchman (1990). Here we describe a more general test.

German

The type frequency hypothesis exploits a correlation: the inflectional form that serves as the default in generalization is also the majority form in the language. The hypothesis assumes a causal relation: frequency in the input to a pattern associator causes greater tendency to generalize. As in many correlations, a causal arrow in the opposite direction can be equally or more plausible. The language that provides the input to the child cannot be treated as an environmental given. It is itself the product of generations of learners and could reflect, rather than shape, their generalization tendencies. In particular, English words may be mostly regular because they are the product of a default generalization process -- a rule -- rather than vice-versa.

Recall that regular suffixation is used for borrowings and derivations from other categories. These happen to be two of the major ways in which a language's vocabulary can grow. If a language started off with a rule for a regular inflection, that rule would automatically get first dibs on these new arrivals, regardless of how many or few words originally underwent regular inflection. If for historical reasons, such as foreign invasions or ecclesiastical and scholastic influences, a language experienced most of its vocabulary growth by borrowings, and also allowed easy conversion of words from one category to another -- both true of English -- then its regular vocabulary could shift from being in the minority to being in the majority, with no change in the psychology of its speakers. Thus under the rule theory, majority status and default status are psychologically independent. If they correlate, it could be because there were historical events that put rule-generated forms in the majority. The crucial prediction, then, is that there should be languages in which the default inflection is not in the majority. The pattern-associator alternative predicts that that should not be possible.

In this paper we introduce cross-linguistic evidence from Modern Standard German which offers such a test. At first glance, the complexity of German seems like a connectionist's dream. Mark Twain wrote (1880/1979, p. 187):

> A person who has not studied German can form no idea of what a perplexing language it is. Surely there is not another language that is so slip-shod and systemless, and so slippery and elusive to the grasp. One is washed about in it, hither and thither, in the most helpless way; and when at last he thinks he has captured a rule which offers firm ground to take a rest amid the general rage and turmoil of the ten parts of speech, he turns over the page and reads, "Let the pupil make careful note of the following exceptions." He runs his eye down and finds that there are more exceptions to the rule than instances of it.

Indeed, proponents of pattern associators sometimes point to German plural formation as an ideal test of the Type Frequency Hypothesis. Bybee (1991:86-87), for example, writes that

> there are many cases of morphological systems in which no one pattern represents the regular pattern. For instance, German Plural formation interacting with the gender system offers several strong patterns and not just one rule. For this reason Köpcke (1988) argues that the situation is best handled by [associative] schemas.

Similarly, MacWhinney and Leinbach (1991:137), discussing Pinker and Prince (1988), claim that

> This contrast between the regular pattern as a formal rule and the minor patterns as fuzzy families of memorized exceptions involves a number of questionable assumptions. Perhaps the most serious problem with this view is the claim that all phonological markings have some preeminent regular pattern. Consider a system such as the marking of plurality on the German noun. German plurals can be formed using *-en*, *-er*, *-s*, *-e* or zero as endings, along with possible vowel ablauting. None of these five possible suffixes is statistically predominant (Köpcke, 1988); none can be characterized as being "the regular ending." In a situation such as this, there is simply no regular pattern at all. Are we to draw some sharp line between English and German speakers by claiming that only the former evidence "rule-governed" behavior?

Note that this argument is circular: MacWhinney and Leinbach define "regular rule" in the descriptive sense of "majority pattern" (irrelevant to the psychological issues under discussion) and then argue against rules on the basis that German has no majority pattern (see Seidenberg, 1992, p. 6, for another example of this circularity). The empirical question is whether the lack of a majority pattern means that German has no rule in the psychological sense of a symbol-manipulation operation that can act as the default generalization. This requires looking at the generalization patterns of speakers, not just counting the types in the language.

We will examine two inflectional systems in German: the participle system, which uses forms similar to the English past tense system, and the even more complex plural system. Each of these systems differs from its English counterpart in the balance of regular and irregular patterns. We argue that the non-dominating participle suffix *-t* and the very infrequent plural suffix *-s* serve as the crucial test cases: despite their low type frequency compared to English, they are generalized in heterogeneous default circumstances, just like their English counterparts.

Experiment 1: German Participles

The structure of the German verb paradigm is illustrated in (17). Each verb has three paradigmatic forms: an infinitive, a preterite, and a participle, which we list in that order (e.g., *kaufen - kaufte - gekauft*). Infinitives consist of a stem and the infinitival suffix *-(e)n*. Preterites or simple pasts are uncommon in informal speech; we do not examine them in this paper. Participles are composed of three grammatical elements. The prefix *ge-* occurs in participles whose stem has primary stress on the first syllable (which is true of most German verbs). Because *ge*-prefixation is independent of the choice of stem pattern and suffix, we will not consider it further. (For detailed analysis see Wiese (1992: Section 4.1.). The second element is the verb root, which sometimes involves changes from the infinitive (most notably ablaut, i.e., vowel change). The third is the suffix *-t* or *-n*.[59]

[59]For many German verbs, the verbal stem is the combination of a prefix and a root. For some "separable prefix" verbs, in the preterite, the prefix stands as a separate independent word. For these verbs, the prefixes are always stressed. In the class of "inseparable prefix verbs," the verb never appears independently of the prefix, and the prefix is always unstressed (except for the prefix miβ- where the prefix varies.)

(17)

Weak verbs:
kaufen - kaufte - gekauft "to buy - bought - (has) bought"
diskutieren - diskutierte - diskutiert "to discuss - discussed - (has) discussed"

Strong verbs:
gehen - ging - gegangen "to go - went - (has) gone"
vertréiben - vertrieb - vertrieben "to expel - expelled - (has) expelled"

Mixed verbs:
rennen - rannte - gerannt "to run - ran - (has) run"
bringen - brachte - gebracht "bring - brought - (has) brought"

There are three classes of German verbs, referred to in traditional grammars as "weak", "strong," and "mixed." In weak verbs, the participle is formed by adding the *-t* suffix to the unchanged stem (along with *ge-* where applicable). For example, the verb *kaufen* "to buy" has the stem *kauf*; the participle is formed by adding *ge-* and *-t*, yielding *gekauft*. *Lernen* ("to learn") inflects the same way: *lernen - lernte - gelernt*. Because the participle form of weak verbs is perfectly predictable, we will refer to weak verbs as "regular."

The participles of strong verbs differ from those of weak verbs in two ways: the stem generally changes, and *-en* is suffixed instead of *-t*.[60] For example, the participle of *gehen* "to go" is *gegangen*.

Finally, there is a small group of high-frequency "mixed" verbs whose simple past tense and participle forms undergo stem changes (as in strong verbs) but which take the same regular suffixation seen in weak verbs (*-te* in the past tense and *-t* in participles). They are thus similar to "weak irregular" verbs in English such as *sleep-slept* and *tell-told*.

The stem changes that the strong and mixed verbs undergo are reminiscent (and historically related to) the stem changes of English irregular verbs. First, in nearly every

---

[60]There is a group of strong verbs (which take *-en*) in which the participle stem does not change, but for all these verbs the preterite stem changes; for example, *sehen - sah - gesehen*.

case, the participle stem generally shares some phonological content with the infinitive stem, e.g., *gehen - ging - gegangen*; *singen - sang - gesungen*, and *schreib - schrieb - geschrieben.* Second, irregular verbs tend to come in clusters. For examploe, *singen - sang - gesungen* strongly resembles *sinken - sank - gesunken* and *trinken - trank - getrunken*; *sehen - sah - gesehen* resembles *lesen - las - gelesen* and *geben - gab - gegeben.* These stem-past and pair-pair similarities suggest that the stems of German strong participles are not just a rote-memorized list of isolates. The stem changes, however, cannot easily be described by rules governed by necessary and sufficient conditions. For example, the *singen* group has exceptions like *beginnen - begann - begonnen*, and the *sehen* group has exceptions like *gehen - ging - gegangen.* Furthermore, German speakers sometimes have graded judgments about the preterite and participle forms of certain verbs, like *backen*, which admits of preterites *buk* or *backte* and participles *gebacken* or *gebackt.* These facts suggest that stems are stored in some kind of associative memory, as in English, which can encourage generalization of these strong patterns to similar new forms.

Wiese (1992) and Wunderlich (1992) have argued that the *-t* of the weak verb is generated by default, just like the English *-ed*; it is the product of a rule that adds the *-t* suffix to verbs that lack a strong form in memory, otherwise leaving the stem of the verb unchanged. Several of the criteria we have used for English apply to the German weak suffix. For example, German speakers would use the *-t* suffix to form the participle of low-frequency verbs like *löten* "to weld" (Table 1: #2), and of novel verbs like *faben* (Table 1: #1), and even of novel verbs that do not rhyme with any existing German verbs, such as *rilken, tuden, teinten, quossen*, and *wauden* (Table 1: #3). The weak suffix can also apply to verbs that are homophonous with strong verbs (Table 1: #4). For example, strong-participled *mahlen - gemahlen* "to grind" is homophonous with weak *malen - gemalt* "to paint"; strong *schaffen - geschaffen* "to create" is homophonous with *schaffen - geschafft* "to work." As in English, many of these homophones involve transitive and intransitive counterparts of the same root (see Note [49]), including *quellen* "overflow," *schmelzen* "melt," *schrecken* "fear/frighten," *schwellen* "swell," *stecken* "stick/poke," and *hängen* "hang." Furthermore, *-t* may be used with verbs that rhyme with strong verbs (Table 1: #5). For example, strong-participled *stehlen - stahl - gestohlen* "to steal" rhymes with the weak *fehlen* "to miss"; strong *saufen - soff - gesoffen* "drink (booze)" rhymes with weak *kaufen* "buy" and *raufen* "wrestle"; strong *blasen - geblasen* "blow" rhymes with weak *rasen - gerast* "rage."

Furthermore, because weak suffix is the default, it applies to words with no available listed past root, such as onomatopoeic verbs (Table 1: #6) like *brummen* "grumble, growl," *flüstern* "whisper," and *klatschen* "clap." It also applies to verbs that are derived from other categories and thus cannot have listed inflected roots. Indeed existing denominal verbs in German (Table 1: #12) are all suffixed with *-t*, e.g., *frühstücken - gefrühstückt* from *Frühstück* "breakfast", *baggern - gebaggert* "to dredge" (from *Bagger* "excavator"), *angeln - geangelt* "to fish", *hausen - gehaust* "to house." This is also true of verbs derived from adjectives, such as *kürzen - gekürzt* "shorten" from *kurz* "short," *säubern - gesäubert* "to clean" from *sauber* "clean," and many others. The effect occurs not just for existing denominal verbs, but for novel ones: for example, the participle of *gorbatschowen*, "to Gorbachev," must be *gegorbatschowt*. Furthermore, it occurs for verbs formed from nouns formed from irregular verbs, that is, verbs that have irregular roots but do not have them in a head position from which the irregular form can percolate (Table 1: #13). For example, the irregular verb *halt - hielt - gehalten* "to hold" can be converted into the noun *Halt* "a hold" which can be used in the compound *Haushalt* "household." The compound can then be turned back into a verb ("to housekeep"), but the irregular forms are unavailable and the regular suffixes apply: *haushalten - haushaltete - gehaushaltet*. (These parallelisms will be summarized in Table 6.)

Wunderlich (1992) and Wiese (1992) express the default status of *-t* in slightly different ways: Wiese proposes that a rule adding the suffix *-n* applies only to "roots," and the rule adding *-t* applies to "stems," and suggests that only irregular participles have a "root" entry. Wunderlich makes the suffixation of *-n* dependent on the presence of a feature [+participle] as part of the base, while suffixation of *-t*, since it is the default, does not require this or any other feature; rather, applying the rule is what supplies that feature. A third possibility is that the strong participles are stored unanalyzed, including the suffix. But in all three accounts, suffixation of *-t* is the default, and suffixation of *-n* depends on word-specific idiosyncratic information.[61] In sum, by the same tests that establish that *-ed* is the default past tense suffix in English, *-t* is the default in German.

Crucially, the numerical differences between German past participles that have regular and irregular forms are very different from those in English. How one quantifies this difference depends on whether one assumes that learning is driven by tokens (the learner processes past tense utterances from parental speech, so more frequently-used verbs are processed more often) or by types (the learner processes past tense entries from the mental dictionary, so each verb is processed the same number of times). See Marcus, et al. (1992, Chapter 5) for extensive discussion. Since Rumelhart & McClelland fed verb types to their model, and because the connectionists' claim about the domination of regular verbs in English vocabulary statistics pertains only to types (regular verbs are actually a minority of tokens; Marcus, et al., 1992), we will count types. Since the total number of types in a language is neither countable (it depends on how many specialized and archaic literatures one samples) nor relevant (presumably only the relatively commonly used verbs are the basis for inflection learning), the most relevant comparison would involve a constant number of relatively common verb types.

---

[61]A similar set of possibilities exist for the *-t* of the mixed verbs: the *-t* may be stored in the lexicon, along with the stem, which, as it is idiosyncratic, must be itself be stored (c.f. Wunderlich, 1992), or the mixed verbs may lack the feature "root," possessed by the strong verbs, that blocks the default *-t* suffixation (Wiese, 1992).

In Ruoff's (1981) sample of 500,000 tokens of spoken German, there are 105,939 verb tokens. The 1000 most frequent types account for 96% of the tokens. Table 3 shows the relevant comparison for the 1000 most frequent verb types, and the comparable estimates from English, based on Francis and Kucera (1982). In both languages, the mean token frequency for regular verbs is far lower than for irregular verbs (Table 1: #2). Crucially, whereas in English regular types are in an 86.4% majority, in German regular types are in a 45% minority. Even adding the mixed verbs to the regular set does not bring the total proportion into the majority. The ratio of regular to irregular tokens is also greater in German than in English. Similar estimates come from Meier's (1964) data on the 1200 most frequent word forms in German, which included 34 participles: 68% were strong, 23% weak, and 9% mixed. (The German statistics come from the spoken language and the English statistics from the written language, and this might be thought to be a problem because the lower-frequency regular verbs are presumably more common in the written than in the spoken language, see Pinker & Prince, 1988; Marcus, et al., 1992. However, the fact that the samples are defined by the 1000 most frequent *types* should minimize this difference.)

Table 3:
Proportions of Regular and Irregular Verbs among the 1000 Most Frequent Verbs

German:

	Strong (Irregular)	Mixed	Weak (Regular)
Ruoff:			
Type Frequency	50% (502)	5% (50)	45% (448)
Token Frequency	47% (47,800)	32% (32,544)	17% (17,289)
Mean Token Freq.	190.4/million	1301.8/million	77.2/million
Meier:			
Type Frequency	68% (23)	9% (3)	23% (8)

English:

	Irregular	Regular
Francis & Kucera:		
Type Frequency	14% (136)	86% (864)
Token Frequency	60% (93,083)	40% (63,158)
Mean Token Freq.	684.4/million	73.1/million

This experiment is designed as an independent test of the claim that modern German speakers use the weak suffix, despite its minority status, in the default circumstance of a verb being derived from a noun. If they do, it would show that generalization of a suffix in default circumstances is not an epiphenomenon of high type frequency; if it is, subjects should inflect novel denominal verbs homophonous with existing irregulars with *-en*. We elicited the subjects' choice of perfect participle forms of novel verbs, using the experimental paradigm of Kim, Pinker, et al. (1991). Items from one condition had contexts that encouraged subjects to interpret the verbs as derived from a noun, such as *bepfeifen* "to put pipes on" (from the noun *Pfeife* "pipe"). In the other, baseline condition, the items had existing verb roots (e.g., *bepfeifen* = "whistle for someone"). The roots were presented in contexts that involved a semantic extension of their usual senses. This was done to rule out the possibility, suggested by Lakoff (1987; see Kim, et al., 1991), that mere semantic extendedness, rather than headlessness or rootlessness per se, encourages regularization (for example, the past tense of *fly out* would be regular because its meaning is not a central sense of *to fly*.)

Method

Subjects. Thirty-two adult native speakers of German from northern Germany were paid for their participation.

Procedure. Subjects were presented with a paper-and-pencil test containing 19 test items and 16 filler items, randomly ordered. Each test item consisted of a context paragraph which used a word either as a noun or a verb, and test sentences in which the word was used as the root of a prefixed verb in its participle form. Thus the subjects should have interpreted the test verb either as a denominal verb, if the first mention had been a noun, or as a semantically extended verb root. Each test item contained two sentences, one with each form of the participle: a regular formed with the *-t* suffix and no stem change, and an irregular form with a stem change and the *-en* suffix. Here are two sample items, with English translations (the test participles are rendered as rough English analogues):

Extended:

(From *blasen* "to blow")

Immer wieder wird Otto Meier dabei erwischt, mit Alkohol am Steuer zu fahren. Jedesmal ist das der gleiche Ablauf. Die Polizei hält ihn an und bittet ihn, in ein Röhrchen zu blasen. Kaum will er hineinblasen, färbt es sich schon in bekannter Weise.

"Over and over, Otto Meier is caught driving drunk. Every time it's the same story. The police stop him and ask him to blow into the Breathalyzer. As soon as he blows, the color changes in the usual way."

Über die Jahre hat Otto Meier schon dreizehn Röhrchen verblasen.

Über die Jahre hat Otto Meier schon dreizehn Röhrchen verblast.

"Over the year Otto has already blown/blowed out 13 breathalyzers."

Denominal:

(From *Pfeife* "pipe")

Die kleinen dreieckigen Pfeifen fuer Yuppies sind bei der Kundschaft gut angekommen. Täglich muß Tabakhandler Meier die Regale auffällen, auf denen die Pfeifen ausgestellt werden. Morgens ist daher immer seine erste Sorge:

"The little triangular pipes for yuppies are a success with the customers. Everyday the tobacconist, Meier, has to fill the cabinets on which the pipes are exhibited. Therefore, his first concern every morning is:"

Sind die Regale auch schon bepfiffen?

Sind die Regale auch schon bepfeift?

"Are the cabinets already been pippen/piped?"

Subjects were asked to rate each test sentence on a 7 point scale, with 7 = "totally natural" and 1 = "totally unnatural." In the instructions to the subjects, the scale was explained using grammatical and ungrammatical examples that did not involve denominals. As in Kim, et al. (1991), subjects were instructed to pay attention to the context sentences, to rate the two test sentences independently (i.e., they were shown, with the help of non-denominal examples, that a verb can have regular and irregular forms that are both natural, or both unnatural), to base their ratings on what sounds most natural to their ears as opposed to what they would guess "sophisticated" usage would consist of, and not to treat the items as examples of jokes, puns, or wordplay.

Each denominal verb had a homophonous counterpart consisting of an extended verb root with the same prefix. Fifteen verbs had inseparable prefixes (*verspinnen, verblasen, verbacken, verbergen, verklingen, verscheren, verschlingen, vertragen, verwiegen, begraben, bescheinen, bereiben, besitzen, befliegen, bepfeifen*). The remaining 4 had separable prefixes (*auslicgen*, *ausweichen*, *abpreisen*, *abringen*). An individual subject saw a given verb either as a semantically extended verb root or as a denominal verb. There were four counterbalanced versions of the questionnaire, such that each subject saw approximately half the verbs as denominal verbs, and half the verbs as semantically extended roots. Each subject received half of the test sentences with the regular participle before the irregular participle, and the other half of the test sentences with the irregular participle before the regular participle; the ordering of regular and irregular participles was reversed for half the subjects.

The 16 filler items contained an existing regular or irregular verb, used in a standard sense, half with separable prefixes and half with inseparable prefixes. Ratings of these items were not analyzed.

Results and Discussion

Figure 1 presents the mean ratings of regular and irregular participle forms of the denominal and extended verbs. As predicted, subjects judged regularly inflected participles as better than irregularly inflected participles when the verb was denominal (means 3.3 vs. 2.1), but preferred irregularly inflected participles to regularly inflected participles when the verb was merely extended (means 5.1 vs. 2.0). The interaction is statistically significant with subjects as the error term, $F(1,28) = 105.83$, $p < .0001$, and with items as the error term, $F(1,18) = 88.69$; $p < .0001$. Sixteen of 19 denominal items were judged as better in the regular versions (2 items were given the same mean ratings), and 18 of 19 extended items were judged as better in the irregular versions (the nineteenth was judged as better as a regular).

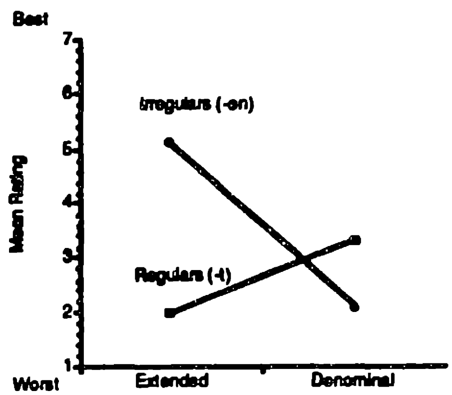


Figure 1. Mean ratings of regular (*-t*) irregular (*-en*) participle forms of novel German verbs that were presented either as semantic extensions of existing German verb roots (Extended) or as derived from German nouns (Denominal).

Thus in German, as in English, denominal verbs are judged as better when regularly inflected, even when homophonous with verb roots that take irregularly inflected participles. Moreover, mere novelty or extendedness, unlike denominal derivation, is not enough to reverse the usual preference for the strong participle forms of these irregular roots.

Clahsen and Rothweiler (1992) presented evidence suggesting that *-t* is treated as a default suffix by German children as well (Table 1: #17). In a longitudinal study of 22 German children (19 were language-impaired, but they behaved similarly to the unimpaired children), they found that the children systematically overapplied *-t* (at rates comparable to those of English-speaking children; Marcus, et al., 1992), but not *-en*: they found 94 instances of overapplied *-t*, but only 4 instances of overapplied *-n*, despite approximately equal numbers of opportunities. Furthermore, there were no examples of a participle in which an irregular stem pattern was extended to a weak verb. The only stem errors they found were 68 cases where the infinitival stem was used in the participle, the regular pattern, rather than the correct irregular ablauted stem. These data further demonstrate that

German speakers generalize regular participle formation more freely than irregular participle formation.

These results demonstrate that when inflecting verbs, German speakers take into account the morphological structure of verbs, and use the *-t* suffix as the default way of forming past participles. Strikingly, verb inflection processes in English and German are qualitatively similar despite quantitative differences in the balance between the number of regular and irregular verbs in the languages' vocabularies.

Experiment 2: German Plurals

Though the German participle system allows some comparison to English, the default suffix in the participle system still applies to nearly half of German verbs. Furthermore, though all the irregular verbs take the suffix *-en*, there are a number of ablaut (stem vowel change) patterns among them, so no single irregular pattern across the whole word is as frequent as *-t* suffixation. The more intricate German plural system provides an even stronger test of the type frequency hypothesis, because no single plural form applies to a large percentage of nouns.

One attempt to account for plural noun formation in German through descriptive (non-generative) rules led to ten rules and nineteen lists of exceptions (Mugdan 1977:87ff.) There are five plural suffixes, *-(e)n*, *-s*, *-e*, *-er* and zero, three of which also allow a variant with an umlaut (Table 4). We will ignore umlaut for the remainder of the paper, because the presence or absence of umlaut within plurals seems to be governed by an independent phonological rule of fronting which applies under specific morphological conditions (Wiese, 1987). For similar reasons ignore the allomorphy between *-en* and *-n* (Wiese, 1992: Section 4.3.1). Orthographic *-e* in suffixes always represents stressless schwa.

Table 4: The plural allomorphs in German

zero (+ umlaut)	der Daumen	die Daumen	"the thumb/thumbs"
	die Mutter	die Mütter	"the mother/mothers"
	der Apfel	die Äpfel	"the apple/apples"
-e (+ umlaut)	der Hund	die Hunde	"the dog/dogs"
	die Kuh	die Kühe	"the cow/cows"
-er (+ umlaut)	das Kind	die Kinder	"the child/children"
	der Wald	die Wälder	"the forest/forests"
	das Huh	die Hühner	"the hen/hens"
-(e)n	die Straβe	die Straβen	"the street/streets"
	die Frau	die Frauen	"the woman/women"
	das Bett	die Betten	"the beds/beds"
-s	das Auto	die Autos	"the car/cars"
	der Park	die Parks	"the park/parks"

To varying degrees, the use of these forms with specific nouns is arbitrary. There exist preferred plural allomorphs according to the gender and/or the morphophonological characteristics of the noun; the list of exceptions, however, is quite long (see, e.g., Twain, 1880/1979; Köpcke, 1988; Mugdan, 1977). For example, masculine and neuter nouns ending with final schwa syllables such as *-er* and *-el* usually form the plural with zero, yet plural forms such as *Bauern* "farmers," *Vettern* "cousins," *Muskeln* "muscles," and *Pantoffeln* "slippers" exist as well. Even families of rhyming words exhibit exceptions, like, *Kind-Kinder, Rind-Rinder,* but *Wind-Winde.*[62]

The *-s* plural suffix is in a decided minority in the German language. We present measures of the frequency of nouns pluralized with *-s* from three different sources. First, Janda (1990) examined a corpus of 600,000 words in taped interviews (Pfeffer, 1964). He

[62]Derivationally suffixed words have completely predictable plural forms. For example, words suffixed with *-ing* takes the *-e* plural; words suffixed with *-er*, *-ler* and *-ner* require zero, and the majority of the feminine derivational suffixes such as *-e*, *-schaft*, *-heit*, *-keit* and *-ung* form the plural with *-(e)n*. The simplest explanation is that the derivational suffixes themselves are just like words, some with their own irregular plural forms; these suffixes may then serve as the heads of the complex words they are part of (see Lieber, 1980; Williams, 1981; Selkirk, 1982; but also Anderson, 1992).

found that of the 200 most common German nouns in this sample (with token frequencies from 100 to 2500 per million), 42% of the types have plurals with *-(e)n*, 35% with *-e*, 12% with zero, 10% with *-er*, and only 1% (*autos* and *hobbys*, 115th and 134th in token frequency) with *-s*. Another five nouns take *-s* as an alternative, nonstandard plural form (such as *Jung-s* rather than *Jung-e* "boys"), raising the estimate to 3.5%.

Second, Janda analyzed a standardized list of cross-linguistically common words (the "Swadesh list," occasionally used in historical linguistics to measure language change.) He found that *-s* was the least common suffix, not appearing at all in the list of 87 nouns. The other suffixes were ranked as follows: *-e*, 39%; *-en*, 28%; zero, 17%; *-er*, 11% (the other 5% allow no plural form).

Third, we conducted our own analyses using the more extensive CELEX database. The CELEX database consists of approximately 381,000 types taken from corpora of 6,000,000 tokens, most from written text of a variety of discourse types (e.g., literature, newspapers, science, etc.), but 600,000 from speech. Celex contains 11,172 noun types that occured at least once in the textual database (some other items appeared only in dictionaries, and hence had frequencies of 0 per 6,000,000). Of these, 6,601 are marked as morphologically complex (i.e., containing a final head that occurs elsewhere). We excluded these, because the German plural is always determined by the final morpheme (see Note [62]); this leaves 4,571 stems. Of these stems, 7.2% (330) take an *-s* plural. The 4,571 stem types correspond to 184,093 tokens in the database. The *-s* plural has a token frequency of 1.9% (3,515) tokens. The proportion of tokens is lower than the proportion of types because, as in English, many words taking *-s* are rare, as reflected in the low average frequency for nouns taking *-s* in both languages (Table 1: #2). This explains the small discrepancy with Janda's estimates, which contain only common words and thus tend to underestimate the proportion of nouns taking *-s*.

The frequency statistics are summarized in Table 5. Each of these measures has drawbacks; Janda's estimates are derived from a relatively small corpus, and the CELEX estimates are based largely on automated analyses that are not yet completely hand-checked. But since they provide converging estimates, we can be confident that the error is small and that frequency of *-s* as measured by types is less than 8%, and by tokens, less than 2%.

Table 5:
Proportions of Nouns taking -*s*-Plurals and Other Plural Forms

German:

	non-*s*-plural	**-*s*-plural**	
Pfeiffer:			
Type frequency	99% (198)	1% (2)	
Swadesh:			
Type frequency	100% (87)	0% (0)	
CELEX:			
Type frequency	93% (4,241)	7% (33)	
Token frequency	98% (180,578)	2% (3,515)	
Mean token frequency		7.1/million	1.8/million

English:

	non-*s-plural*	**-*s*-plural**	
Francis & Kucera:			
Type Frequency	0.4% (24)	99.6% (6,123)	
Token Frequency	5% (2,658)	95% (51,399)	
Mean token frequency		110.8/million	8.4/million

But among the plural allomorphs, *-s* is special in several ways -- exactly the ways that make up the default circumstances of inflection. This may have been first pointed out by Van Dam, 1940, who called *-s* the *Notpluralendung* "emergency plural ending" (also see Wiese 1986; Bornschein and Butt, 1987; Janda 1990). First, the use of *-s* is morphophonologically free: *-s* appears when the phonological environment does not permit any other plural allomorph (Table 1: #3). The morphophonological space of German noun plurals varies on several dimensions, including gender, syllable structure, and rhyme structure. Irregular suffixes are restricted to particular regions. For example *-er* applies predominantly to neuter nouns, and never to feminines. But noncanonical words may fit in regions outside of the phonological space in which the only suffix that treads is *-s*. This is especially obvious in the case of unassimilated borrowings (Table 1: #9). For example *Café*, which has non-canonical stress, takes the *-s* plural, as does unusual-sounding *Kiosk*. Köpcke (1988:325) found that of 182 recent German borrowings, about half were formed with *-s*. However, as in other languages, some historical borrowings can become assimilated into the canonical template for roots; in such cases an irregular can be applied, such as *Computer-Computer, Firma-Firmen, Manuskript-Manuskripte*.

Furthermore, *-s* can even appear with stems that rhyme with existing irregular nouns (Table 1: #5). Bornschein and Butt (1987, p. 142-143; translation ours) argue that

> .... it is important to state that the *s*-plural mainly occurs independent of such factors, i.e., with words whose phonological features in other cases allow other plural markers: *Reelings* "railings" but *Ringe* "rings," *Schecks* "cheques" but *Flecken* "spots," *Labels* but *Kabel* "cables," *Tiefs* "lows" but *Briefe* "letters," *Riffs* "reefs" but *Kniffe* "tricks," etc. Thus, the suffix *-s* is the only plural that can appear in any part of the morpophonological space.

Moreover, there is a wide variety of special grammatical circumstances in which *-s* plural trumps all other plurals, regardless of phonology. In German, as in English (Table 1: #8), pluralized nouns based on names homophonous with irregular nouns must take regular inflection, hence *Manns/*Männ/*Männer, Wieses/*Wiesen*. This occurs not only with semantically opaque surnames, but with product names (Table 1: #14) that have salient canonical roots but are headless. For example, the car model Opel *Kadett* would be pluralized as *Kadetts*, despite the fact that the common noun *Kadett* "cadet" forms its plural as *Kadetten* (similarly, we find Volkswagen *Golfs und Jettas*, based on the names of winds).[63] Similarly, eponymous movie or play titles take the *-s* plural: *Faust - Fäuste* "fists" versus *Faust - Fausts* "productions of the play *Faust.*" Furthermore, *-s* is used as the exclusive plural for onomatopoeic nouns (e.g., *Kuckucks* "cuckoo," *Wauwaus* "dogs"; Table 1: #6), quoted nouns (e.g., *Nach korrekturlesung für sexistische wortwahl fand ich drei "Mann"s auf seite 1*; Table 1: #7), nouns based on other categories like conjunctions (e.g., *wenns and abers* "ifs and buts"; Table 1: #12) and verb phrases (e.g., *Vergißmeinnichts* "forget-me-nots" and *Rührmichnichtans* "touch-me-nots"; Table 1: #16), acronyms (e.g., *Gmbh.-Gmbhs* "corporations"; Table 1: #11), and truncations (e.g., *Wessi - Wessis* from *Westdeutsche* "West Germans," *Sozi - Sozis* from *Sozialist*; Table 1: #10).

A corroboration of the special status of *-s* comes from a circumstance in which it *cannot* occur. In German, as in English, regular plurals are generally excluded from compounds, though plurals with the other forms can appear inside them (Table 2: #6). For example, the compounds containing irregular plurals with *-e*, *-en*, and *-er* in the following examples are acceptable, whereas the compounds containing regular plurals with *-s* are not. (For further discussion, see Clahsen et al., 1993; Wiese, 1992, Chapter 5.)[64]

(18)

Professor-*en*-kränzchen "professors' circle"
Frau-*en*-laden "women's center"
Schwein-*e*-stall "pigsty"
Gäns-*e*-braten "roast goose"
Hühn-*er*-ei "hen egg"
Büch-*er*-regal "bookshelf"
Sozialist-*en*-treffen "socialists' meeting"

*Sozi-*s*-treffen "socialists' meeting" (clipped form of Sozialist)
*Auto-*s*-berg "cars heap"

---

[63]Note that Citroen *Ente* would be pluralized *Enten*, because the cars are reminiscent of ducks and thus the plural name may be construed as based directly on the plural noun. Similarly the Volkswagen *Käfer* "beetle" is pluralized as *Käfer*, presumably because of the car's beetle-like shape. See Note 12.

[64]An *-s* suffix is sometimes permitted in compounds (e.g., the *-s* in *Wirtschaftskrise* "economic crisis"), but only as a linking element (*Fugenelement*), similar to *huntsman* and *bondsman* in English. See Clahsen et al. (1993), Wiese (1992, Chapter 5), and Pinker & Prince (1988).

Finally, some German children overregularize with the *-s* plural (Table 1: #17; Park, 1978; Veit, 1986; Schaner-Wolles, 1988; Clahsen & Marcus, 1992; Clahsen, et al., 1993).[65] This fact provides further evidence that *-s* is the default.

In fact, the only nonparallelism between English and German *-s* we have discovered occurs in Bahuvrihi compounds, which in German inherit the plural form of the rightmost morpheme (e.g., *Stillleben* "still lives" from *leben* "lives"). Apparently, percolation of the referent and of the morphological forms of a noun in German do not always follow the same path: the latter obeys the right-hand head principle even when the former does not.

In sum, the *-s* plural displays a range of applicability and non-applicability that resembles that of the English *-s* plural in at least 13 different ways. These properties leads to the characterization of *-s* as the "default plural" or "emergency plural." This, we suggest, bespeaks the operation of a rule referring to a symbol for an entire category that applies unless it is specifically blocked by a competing lexically stored form.

Partly because these criteria have not invariably been applied in the literature, not all German linguists have treated *-s* as the default plural. For example Wurzel (1990), while in general an advocate of rules, does not adopt Wiese's (1986, 1988) suggestion that *-s* is the default plural marker in German, but claims instead that a number of inflectional rules operate to give nouns their unmarked plural form (e.g., monosyllabic feminines receive *-n*, bisyllabic vowel-final nouns receive *-s*). All the rules have equal status in the grammar. This position is reminiscent of the treatment of descriptive handbooks and dictionaries (e.g., Helbig & Buscha, 1987; Wahrig, 1986), which list a number of inflectional types for nouns in German, and assign each noun to one of the types.

Coming from the other direction, Köpcke (1988) has argued that German plurals are

---

[65]Many children also overextend *-en*, but interestingly, such children also omit it from compounds where the adult language permits it, a signature of a regular default (Table 2: #6). This suggests that such children have temporarily misinterpreted *-en* as the default.

assigned by matching the stem to associative schemas. The schemas are envisioned as capturing the similarities among the existing nouns that take a plural suffix; nothing has a special status as the default. Thus Köpcke's schemas are an example of a uniform model of model of morphology of the kind advocated by Bybee and by MacWhinney and Leinbach. Köpcke asked German adults to pluralize 50 novel noun stems which varied by gender (masculine, feminine, and neuter) and syllabic structure (suffixed, schwa-final, vowel-final, pseudosuffixed, or monosyllabic). He found that each type of word has a preferred suffix. For instance, feminine schwa-final nouns nearly always took *-(e)n*, and monosyllabic nouns tended to take *-e* if neuter or masculine, *-(e)n* if feminine. But Köpcke's experiment did not test whether there is a default pluralization process in German. Each noun was presented in isolation, so subjects presumably treated them as canonical roots. Any model that acknowledges that the memory for word roots fosters analogies would predict that novel roots may take the suffix of similar existing roots. To test whether there is a default plural rule in German, one needs to examine the pluralization of novel nouns in various circumstances in which access to memory for roots is ruled out.

In Experiment Two, we test pluralization in heterogeneous circumstances held together only as the default. The design crosses three factors: Regularity (Regular vs Irregular), Rhyme (Rhyme vs Non-Rhyme), and Root (Root vs Name vs Borrowing).

The Regularity factor classifies *-s* as the Regular plural suffix, for reasons just discussed. In contrast, *-e* and *-er* are clearly Irregular. Although a case could be made that *-en* is sometimes generated by a rule that operates on specified subcategories of nouns (e.g., feminine nouns ending in a schwa, which all take *-en*), we treat it as Irregular here, because it clearly is not rule-generated in the case of monosyllabic nouns, the kind of item we use in the experiment. Moreover, Wiese (1986) provides an alternative explanation for the subcategories for which *-en* is predictable: if the schwa on feminine nouns is itself a derivational suffix, it can select *-en* just like any other irregular morpheme (see Note [62]).

The Rhyme factor tests the hypothesis that novel roots are likely to receive irregular inflection if they are similar to existing words; otherwise, the default inflection is available to inflect them. Novel nouns were either "Rhymes," words that rhyme with existing German irregular nouns (and that do not rhyme with regular nouns), and "Non-Rhymes," words that do not rhyme with existing German nouns. If the *-s* is applied by a default rule, German speakers should judge the *-s* plural as better for Non-Rhymes than for Rhymes, because Non-Rhymes are less likely to evoke clusters of irregular roots in memory, allowing them to slip directly into the default process. Judgments of irregular suffixes should show the opposite pattern: they should sound poorer with the unusual-sounding non-rhymes, which fail to evoke the relevant analogy-fostering existing irregulars. In contrast, a single pattern associator appears to predict that all suffixes should be weaker for unusual-sounding words than for canonical-sounding words.

The morphophonological space of German noun plurals is vast, so we sampled only a small region, monosyllabic nouns. These items should be particularly drawn to the irregular clusters of *-e*, and to a lesser extent *-er*, since those suffixes are the ones most commonly used to inflect monosyllables; monosyllables are rarely inflected with *-s*. Hence this choice works against our hypothesis.

The Root Factor has three levels: Roots, Names, and Borrowings. A third of the words were presented as Roots, that is, as normal German words in a neutral context. All of these words are grammatically eligible either for regular or irregular inflection, and the choice should be determined largely by Rhyme: though all the roots could be analogized to the irregular patterns, since they are all monosyllables, the analogy clearly should be stronger for Rhymes, which should thus take irregular plurals to a greater extent (and regulars to a lesser extent) than Non-Rhymes would.

Another third were presented as Names, a circumstance that should elicit the regular or default plural form, *-s*. Since information about the phonological structure of roots is systematically withheld from the representation of the entire word, in principle Rhyme should make no difference; all names should be suffixed by the default process.

The remaining third were presented in a context that suggested that they were Borrowed from a foreign language. As in the case of English, German contains cues as to the native versus borrowed status of morphemes. For example, Latinate suffixes can be distinguished from Germanic ones because only the former may bear stress. But as discussed earlier, borrowings can sometimes be assimilated to root status, especially if their phonological patterns fit the canonical template for the language. The prediction is that when speakers assimilate a borrowing, it should behave like a Root and hence take irregular inflection when similar to existing irregulars, regular inflection otherwise. But when speakers treat a Borrowing as a borrowing, it should take regular inflection across the board (for the same reason that Names do). The Rhyme factor thus enters into the predictions in two ways, because while all the novel forms, as monosyllables, should have some likelihood of being assimilated, the Rhymes, which resemble existing forms more than the Non-Rhymes do, should have an even greater likelihood. In addition, once assimilated, Rhymes should be more likely to elicit irregular suffixes by analogy to existing forms than Non-Rhymes. The overall predictions for the Borrowings, then, are straightforward, if a bit complex: Among the Borrowed Rhymes (assimilable and analogizable) there should be a preference for the irregular forms over the regular ones, though not as big a preference as seen among the Roots, since the question of assimilation does not arise for them. Among the Non-Rhymes, there should be a preference for the regular forms over the irregular ones, though not as big a preference as seen among the Names, since some of the Borrowed Non-Rhymes might still be assimilated and elicit analogization, generally impossible for the Names. In other words, for the Borrowings, the interaction between Suffix and Rhyme should fall somewhere in between what is found for the Roots and what is found for the Names.

The predictions of the theory that all inflection is computed in a single pattern associator are quite different, because pattern associators neither easily generalize low-frequency suffixes, nor unite the different default circumstances (phonological and derivational) as defaults. Associative models that rely solely on phonological information (e.g., Rumelhart and McClelland, 1986; Plunkett and Marchman, 1991; Daugherty and Seidenberg, 1992) must predict that Root has no effect, since it cannot even be represented in the input. These models predict that *-s* should be eschewed across the board: driven only by phonological similarity, the models should always prefer the more common *-e*, *-en*, and *-er* plural forms to *-s*, even for Non-Rhymes, since there is no reason that very rare *-s* would scoop up the words that have lower similarities to existing irregulars and no similarity to existing regulars. Precise predictions about a hypothetical pattern associator model that would somehow represent root and head structure in the input (and hence the difference among Roots, Names, and Borrowings) must remain somewhat conjectural. Assuming that such a model could be given a plausible "Roothood" feature node or its surrogate, it probably would learn from the few *-s* forms in its input (presumably all non-roots) that this feature predicts an *-s* plural. But given the rarity of *-s* plurals both typewise and tokenwise, it would have no way of knowing that *-s* is also more applicable to unusual roots, nor that for Names, *-s* overrides any effects of phonological similarity and applies to all pluralized names, Rhyme and Non-Rhyme, equally strongly.

Method

Subjects. Forty adult subjects were recruited from Northern Germany.

Procedure. Subjects were presented with a paper-and-pencil test requiring them to judge plurals of novel words. There were four versions, each containing 24 items. Each item contained a novel word presented as a singular form in a context sentence, followed by a set of test sentences containing each of the possible plural forms for the novel word. Subjects were asked to rate each sentence on a scale from 1 "perfectly natural" to 5 "perfectly unnatural." The endpoints were anchored this way so as to correspond to the sequence of grade scores familiar in German schools. To make the data commensurable to those from other experiments, we subtracted each rating from 6, so that higher numbers would correspond to ratings of greater naturalness. The subjects were asked to rate each item in terms of how "normal" or "good-sounding" as opposed to how "funny" or "wrong" they were. They were told not to pay attention to the orthography of the words, only their sounds, that nouns could have any number of natural or unnatural plural forms, and that there were no right or wrong answers; only their personal evaluations were of interest.

The 24 items were divided into 2 (Rhyme/Non-Rhyme) x 3 (Root/Name/Borrowing) = 6 conditions, with 4 novel words appearing in each condition. Rhymes were selected to rhyme with large clusters of irregular German nouns. For example *Pund* was used on analogy to *Hund-Hunde*, *Pfund-Pfunde*, *Grund-Gründe*, and so on. The Rhyming items consisted of *mur, bral, raun, nuhl, pisch, pund, vag, kach, spert, pind, spand, klot.*[66] Non-rhymes were created using a table from Seiler (1970:417) which contained lists of permissible and non-permissible combinations of German onsets and codas. Possible but non-existing combinations were selected: *Bnaupf, Bneik, Bnöhk, Fnähf, Fneik, Fnöhk, Plaupf, Pleik, Pnöhf, pröng*, and *snauk*.

Items were presented as Roots by introducing them as novel but otherwise ordinary German nouns. For example, one lead sentence was

Wuβten Sie, daβ dieses kleine Dingsbums eine KLOT ist?

"Do you know that this little thingamajig is a KLOT?"

Subjects then were faced with each of the following continuations:

Es gibt 3 KLOT in meiner Werkzeugkiste.

Es gibt 3 KLOTE in meiner Werkzeugkiste.

---

[66]We inadvertently selected one item (*Bral*) which rhymed with the low-frequency word *Schal* "scarf," which permits either the plural *-s* or the plural *-e*. This is unlikely to be a significant problem, because *Schal* rhymes with five other words, *Tal-Täler, Mal-Male, Wahl-Wahlen*, and *Zahl-Zahlen*, all more frequent than *Schal*.

Es gibt 3 KLOTER in meiner Werkzeugkiste.

Es gibt 3 KLOTEN in meiner Werkzeugkiste.

Es gibt 3 KLOTS in meiner Werkzeugkiste.

"There are 3 KLOTS in my toolbox"

Items were presented as Names by introducing them as the surname of each of a set of people. For example, one lead sentence was

Mein Freund Hans KACH und seine Frau Helga KACH sind ein bißchen komisch.

"My friend Hans KACH and his wife Helga Kach are a bit strange."

followed by five test sentences, one per plural form, like

Die KACH versuchen immer, ihre Schuhe anzuziehen, bevor sie die Socken anhaben.

"The KACHS always try to put on their shoes before they put on their socks."

Items were presented as Borrowings by introducing them as foreign words for various objects (though never as proper names).[67] An example lead sentence was

Haben Sie schon Mal den PLEIK in dem neuen russischen Restaurant gegessen?

"Have you tried the pleik in the new Russian restaurant?"

followed by five test sentences like

Man braucht 3 PLEIK, um satt zu werden.

"You need 3 pleiks in order to be satisfied."

Except for personal names, which were unmarked for gender (following the common German pattern), half the items in each questionnaire were masculine, half feminine; across

---

[67]Due to experimenter stupidity, one sentence context for borrowed items, *Der Computerladen...*, appeared twice, with different test words, in every questionnaire.

subjects, each item appeared an equal number of times as masculine and as feminine. (We included no neuter nouns, because they pluralize much like masculine nouns.) The questionnaires were presented in one of two orders, counterbalanced across subjects; one order was assembled at random, the other was its mirror image. In addition, there were two sets of questionnaires, one given to 16 of the subjects, the other to 24. In one set, each word was assigned at random to either the Root, Name, or Borrowing condition. In the other, the questionnaires were counterbalanced across subjects so that each word appeared an equal number of times as a Root, Name, or Borrowing. Data from the two sets of questionnaires were nearly identical, and the difference between sets did not statistically interact with any other factor.

In conducting the analyses, we were maximally charitable to the pattern associator hypothesis by comparing each subject's rating of the *-s*-suffixed form of a given noun to his or her *highest* rating among all the irregularly-suffixed forms of that noun. For example, if a subject rated *Klote* as 4, *Kloter* as 3, and *Klots* and *Kloten* as 2, we would use 4 as the rating of the "Irregular" form for that item for that subject. The comparison between best Irregular and *-s* was treated as a within-subjects factor, Regularity, in the analyses of variance to be presented below. (Alternative measures, such as the mean rating of the irregular forms, would systematically underestimate the strength of irregulars in default contexts and thus help our hypotheses.) The zero-suffix or no-change forms were not analyzed because subjects actually saw these forms in the context sentences, possibly biasing them to rate it higher, and because subjects might interpret the no-change form as reflecting unwillingness to pluralize the noun at all, as opposed to reflecting their choice of how to pluralize it given that it must be pluralized. In any case, adding back the zero forms has little effect on the results.

Results and Discussion

Mean ratings of the *-s* plural and of the best irregular plural forms of Roots are presented in Figure 2a. Overall, the ratings were better for the best irregular. Presumably this reflects the strong pull exerted by existing German monosyllabic noun roots, the vast majority of which have irregular plural forms. Irregular forms were judged as better in the Rhyme condition than in the Nonrhyme condition (4.3 versus 3.9), whereas *-s* suffixed forms were judged as worse in the Rhyme condition than in the NonRhyme condition (3.3 versus 3.6). This interaction was significant by subjects $F(1,39) = 20.08$, $p < 0.001$, and by items $F(1,22) = 4.63$, $p < 0.05$. For Non-Rhymes, the difference between the ratings of the Regular and Irregular is not significant. The decline in the response of the irregular items can be explained as a standard associative generalization gradient. The improvement of *-s* across this comparison, in contrast, suggests that it is produced as a default (lack of similarity to irregulars is not a sufficient condition for generalization of a regular suffix in a pattern associator; see Prasada and Pinker, 1992, for a demonstration). Though *-s* was not rated better than the best irregular form on average, it was rated better in 31% of the subjects' ratings of the Non-Rhymes, and equal to the best irregular in an additional 24%.

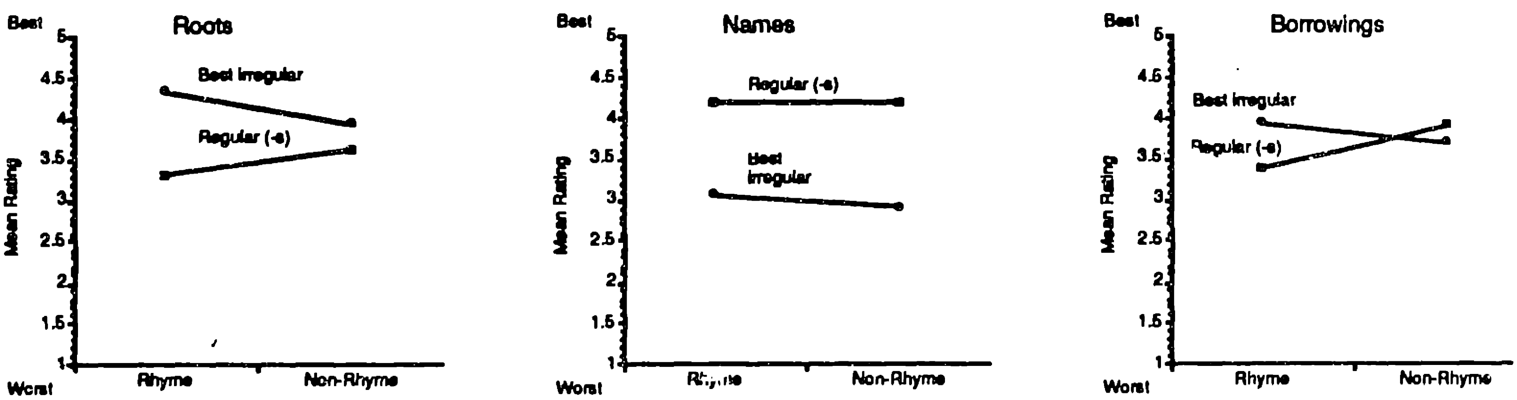


Figure 2. Mean ratings of the regular plural form (*-s*) and irregular plural form (highest rating for that subject and item among *-en, -er, -e*, with and without umlaut) of novel German nouns. Nouns either rhymed or did not rhyme with existing German noun roots, and were presented either as roots (a), derivations from names (b), or borrowings from foreign languages (c).

Mean ratings of the plurals of Names are presented in Figure 2b. Regular (*-s*) plurals were rated as *better* than the best Irregular plural (mean 4.2 versus 3.0; the difference is significant by subjects, $F(1,39) = 39.89$, $p < .001$, and by items, $F(1,22) = 26.11$, $p < .001$). This is exactly the opposite of how Roots were rated. Just as strikingly, *-s* was preferred, and irregular plurals dispreferred, to the exact same extent among Rhymes and Non-Rhymes, as predicted by the rule hypothesis. Specifically, subjects gave identical mean ratings for the Regular plural forms of Rhymes and Non-Rhymes (mean 4.2), and gave very close ratings to their most preferred Irregular plural form of Rhymes and Non-Rhymes (3.1 versus 2.9, a difference which is not significant by subjects $F(1,39) = 1.39$, or by items $F(1,22) < 1$). The contrast between the similarity-sensitivity of Roots and insensitivity of Names can be tested in the three way interaction between Rhyme, Root, and Regularity (Root versus Name); this interaction is significant, by subjects, $F(1,39) = 5.70$, $p < .05$, though not by items, $F(1,22) < 1$.

Mean ratings of the *-s* plural and of the best irregular plural form of Borrowings are presented in Figure 2c. As in the case of Roots, the best Irregular plural form was judged as better for Rhymes than for Non-Rhymes (4.0 versus 3.7), while the regular plural forms were judged better for Non-Rhymes than for Rhymes (3.9 versus 3.4). This interaction is significant by subjects, $F(1,39) = 10.67$, $p < .002$, and marginally significant by items $F(1,22) = 3.68$, $p < .07$.

Note, too, that in comparison with Roots, Borrowed nouns triggered higher mean ratings for Regular plural forms (3.6 versus 3.5), and lower ratings for the best Irregular plural form (3.8 versus 4.1). This is reflected in a significant interaction between Root (Root vs Borrowing) and Regularity, $F(1,39) = 17.10$, $p < .001$ by subjects, $F(1,22) = 8.25$, $p < .01$, by items. As predicted, Rhyme had similar effects in the two conditions.

Finally, we can test another prediction distinguishing the hypotheses. According to the Pattern Associator hypothesis, the various plural forms lie on a continuum from most to least productive in default circumstances, reflecting their type frequencies from most to least frequent. The more frequent and productive a suffix, the better it should sound in nonroots relative to roots. According to the Rule Hypothesis, all the irregular forms are grammatically equivalent, and should suffer to the same degree when put in non-root, default circumstances, where the regular *-s* form alone applies. The German plural forms should display this continuum, because according to the frequency statistics cited above, the ratio of frequencies between the most frequent and least frequent irregular suffix is even greater than the ratio of frequencies between the least frequent irregular suffix and regular *-s*. Figures 3a and 3b illustrate that while the effect of Root is in the opposite different for the irregular suffixes taken together and the *-s* suffix, it is the same for each of the irregular suffixes when they are compared to one another. A quantitative test of this prediction is that the judgments of the *-s* suffix should be negatively correlated with the judgments of the best irregular suffix. By calculating a correlation within each subject between rating of the best irregular form of each item and of the *-s*-plural of that item, and transforming the resulting *r*s to *z*-scores, we get a mean z of -.25 (equiv to a an r of -.24), which is significantly different across subjects from zero (t = -4.16, $p < .001$). This is not because subjects tried to exaggerate the spread among their ratings of different suffixes for a given word: ratings of the different irregular suffixes were positively, not negatively, correlated with each other (all mean pairwise correlations between *-e*, *-er*, and *-en* over subjects were positive and significant or marginally significant). This further suggests that the irregular forms were being treated as qualitatively distinct from the *-s* suffix.

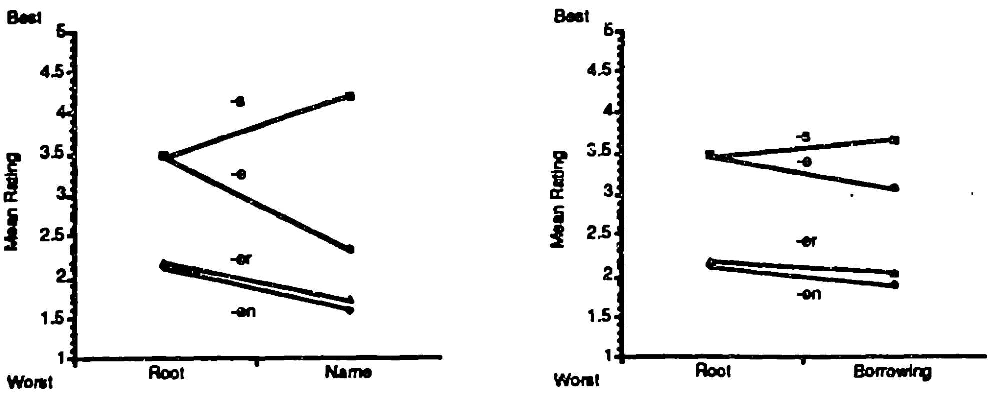


Figure 3. Mean ratings of the regular plural form (*-s*) and irregular plural forms (non-umlauted *-e, -er, -en*) of novel German nouns, contrasting nouns presented as roots and as derivations from names (a), and contrasting nouns presented as roots and as borrowings from foreign languages (b).

In sum, German speakers generalize the *-s* plural and the irregular plural forms in qualitatively different ways. If a novel noun is learned as a root, or as a borrowing easily assimilable to roots, its plural can be formed by analogy to the plural forms of similar existing irregular nouns. The suffix *-s* is used elsewhere: for the more unusual-sounding roots, for unassimilated borrowings, and for names with both usual and unusual sounds. The heterogeneity and rarity of these circumstances argue that the plural *-s* suffix applies not as an association separately acquired for each of these combinations of circumstances, but whenever the combination allowing memory-based generalization (similar root in head position) does not apply; that is, elsewhere, as the last resort, emergency, or default. The fact that *-s* serves as a default even though it is rare argues against the hypothesis that default application of a suffix is an epiphenomenon of its having been experienced with a large number of words. Instead, defaultness appears to be a consequence of the suffixation process accessing a mental symbol for a grammatical category, and hence applying indiscriminately to any word that such a symbol refers to unless specifically blocked.

General Discussion

In this paper, we have set stringent criteria for attributing a linguistic process to a mental symbol-processing operation, such as concatenating a suffix to a variable. The mere existence of common patterns and productive generalization to novel forms is not enough: English irregular verbs display both, and these phenomena must be attributed at least in part to the kinds of analogies that associative memories, such as connectionist pattern associators, are prone to making. English regular verbs, in contrast, afford a kind of generalization that is not dependent on analogies to stored patterns. Exactly the opposite: English regular inflection applies in all and exactly the circumstances where analogies fail, for any reason, whether it be because there are no similar relevant forms to support the analogy, or because access to remembered lexical information is cut off in a word whose derivational structure is designed to prevent the usual transmission of information from the lexicon. We listed 20 heterogeneous circumstances, many quite rare, in which such lexical analogies fail but regular suffixation applies. No existing pattern associator theory can explain this range of generalization, but it coincides exactly with the range of a symbol for a grammatical category like verb or noun, and hence supports the hypothesis that the mind contains rules manipulating such symbols.

The central experimental finding of the paper is that this default application of a linguistic pattern does not depend on the pattern constituting a majority of the learner's experience, the principal hope for pattern associator modelers that future models might duplicate such generalizations. In German, the *-t* participle and *-s* plural, despite much lower type frequency than their English counterparts, are clearly applied as defaults, by the same diverse criteria used to show default inflection in English. This remarkable parallelism is summarized in Table 6 (cf. Table 1).

Table 6:
Circumstances in Which German *-s* and *-t* Suffixes
Behave Like the English Regular *-s* and *-ed* Suffixes

Circumstance	Kind of Word	Example
Lack of Entry or Similar Entries in Memory:		
1. No root entry	Novel words	*gefabt*
2. Weak entry	Low-frequency words	*gelötet*
3. No similar entries	Unusual-sounding words	*gequosst, Fnöhks*
Competing Entries or Similar Entries in Memory:		
4. Competing root entry	Homophones	*mahlen-gemahlen/malen-gemalt*
5. Competing similar root entries	Rhymes	*stehlen-gestohlen/fehlen-gefehlt*
Entry is Not a Canonical Root:		
6. Rendering of sound	Onomatopoeia	*gebrummt, Wauwaus*
7. Mention versus use	Quotations	*drei "Mann"s*
8. Opaque name	Surnames	*Thomas Manns*
9. Foreign language	Unassimilated borrowings	*Cafés, Kiosks*
10. Distortion of root	Truncations	*Sozis, Wessis*
11. Artifical	Acronyms	*Gmbhs, BMWs*
Root Cannot Be Marked for Inflectional Feature:		
12. Derivation from Different Category	Denominal verbs	*geangelt, bepfeift*
	Deadjectival verbs	*gekürst, gesäubert*
	Nominalized conjunctions	*Wenns, Abers*
Features Cannot Percolate from Root to Whole Word (Exocentrism or Headlessness):		
13. Derivation via different category	Denominal nominalized verbs	*gehaushaltet*
14. Derivation via name	Eponyms	*Fausts*
	Products	*Kadetts, Golfs*
15. Referent different from root	Bahuvrihi compounds	(not regularized)
16. Lexicalization of a phrase	Nominalized VPs	*Vergißmeinnichts*
Memory Failures:		
17. Children	Overregularizations	*gesingt, Manns*
18. Normal speech errors	Overregularizations	?
19. Alzheimer's Disease	Overregularizations	?
20. Williams Syndrome	Overregularizations	?

We can rule out several alternative explanations for the data. One springs from the fact that English and German are historically related sister languages. Perhaps, then, the similarities between the behavior of German *-t* and English *-ed*, and between German *-s* and English *-s*, are due to inheritance from their common ancestors in Proto-Germanic, not to a common psychology of their modern speakers. Such an alternative begs key questions, such as how the speakers of Proto-Germanic commanded default affixes if not by a rule, and how default application could survive in modern German despite the extremely low frequency of relevant exemplars. But the main problem is that the mode of inheritance -- a generation-to-generation relearning of the individual circumstances in which the *-t* and *-s* suffixes do and don't apply -- is psychologically implausible. Many of the circumstances that trigger regularization, such as pluralized acronyms or quotations homophonous with irregular nouns, are far too contrived and infrequent to have depended on parallel, 2000-year old unbroken chains of speakers of both languages all using such forms in the presence of their children. For example, Kim, Marcus, et al., 1993 looked for exemplars of one of the circumstances, denominal homophones of irregulars, in 7,500 parental utterances addressed to four children, and found none. Furthermore, one of the key tests for rule-generation, *inability* of regular plurals to appear inside lexical compounds, is demonstrably independent of learning. Children do not have to learn *not* to say *rats-eater*. Nor could they, for they do not produce the error (Gordon, 1985), would not get corrected if they did (Marcus, 1993), and do not hear any plurals, regular or irregular, inside compounds in their parents' speech (Gordon, 1985).

Another possible alternative explanation based on history is that the dominance of *-s* in German was the result of lateral borrowings from English or French, rather than vertical inheritance from Proto-Germanic. Again the question arises of how such borrowed default affixation could have caught on and been maintained given that German speakers have never heard regular words as the majority form. In any case, borrowing of inflected words cannot be the explanation for the current status of the default plural. Köpcke (1988:326) notes that

> Even before the influx of French and then English nouns with *s*-plurals into German a small number of native *s*-plurals already existed in the language (cf. Öhmann 1961-2). Thus *-s* was a low frequency but already recognizable plural morpheme in German. In comparison, other foreign plurals lacking a basis in the native German lexicon have made no inroads whatever. Foreign plural markers

such as Greek *-ta* as in *Thema-Themata*, Hebrew *-im* as in *Cherub-Cherubim*, or Italian *-i* in *Tempo-Tempi* are limited to the small number of stems with which they were borrowed and almost inevitably have a primary or secondary germanicized plural, *Themen, Cherubinen*, and *Tempos*.

Moreover, Modern Arabic is historically unrelated to Germanic, but in its plural system, type frequency and default generalizability can be dissociated (McCarthy and Prince, 1990), just as we have shown for German. The Arabic "broken plural" is the most common set of forms, but is limited to various classes of similar canonically-shaped nouns. The "sound-plural," in contrast, behaves like a default in that it applies indiscriminately to all noncanonical forms, such as proper names, transparently derived nouns such as deverbals and dimunitives, noncanonical or unassimilated borrowings, and the names of letters of the alphabet, which are mostly noncanonical. These forms together, though, constitute a minority of Arabic nouns. In general, the rule theory predicts that default generalization in the absence of high type frequency can appear in any language.

Indeed, the common ancestor of English and German may have been such a language. In Proto-Germanic, most past tenses were formed by an ablaut process that has left fossils in the form of strong verbs. But it also had the weak suffix (the ancestor of *-ed* and *-t*), and used it for derived forms and borrowings (Pyles and Algeo, 1982). Similarly, Janda (1990:148) notes that Old English resembles German in having a multitude of plural forms (at least nine), including the ancestor of *-s*. Modern German remained in the state where *-t* and *-s* served a minority of words; in English, they attained majority status.

Whence the differences between German and English? Two peculiarities of the history of English are that it borrowed many of its verbs from French and Latin (perhaps 60%), and that it invented many other verbs by deriving them from nouns (perhaps another 20%), usually with no overt affixation (Prasada & Pinker, 1983). Borrowings and denominals are two of the nonroot circumstances that require affixation by default --

exactly where the weak suffix did, and still does, apply. Regular verbs are in the majority in English because of the historical accidents that defined the chief growth areas for its verb vocabulary over the centuries. The psychological process applying the regular suffix need not have changed over the millenia as regular verbs grew from minority to majority status.

Janda (1990) goes so far as to predict that "the eventual emergence of *-s* as the dominant plural-ending of NHG seems extremely likely" (1990:149), a conjecture supported by Köpcke's analysis of recent loanwords. Perhaps in several hundred years, as German borrows words from the global community, the *-s* plural will become one of the more frequent affixes. Crucially, although Old English, Modern English, Modern German, and Future German differ in the balance of regular and irregular lexical items in their vocabularies, the speakers of all these languages would appear to have identical psychological processes generating the affixes.

Anyone who doubts that connectionist pattern associator models do not naturally compute grammatical defaults has only to look at Hare and Elman's (1992:267) prediction about historical change: "any mapping with low type and token frequency will be difficult to learn and is likely to be lost." This is exactly the opposite of what has happened in English and German. According to the rule hypothesis, this is no paradox, for the connectionists' interpretation of the correlation between regularity and type frequency in Modern English puts the causal arrow backwards. High type frequency does not cause default behavior; default behavior can cause high type frequency.

Given that the learning of default affixes cannot depend on high frequency, what does it depend on? It is surely not overtly learned in school or from prescriptive grammars. Bornschein and Butt (1987:135-136; translation ours) provide a historical sketch of learned opinion of the *-s* plural.

> The noun category which forms the plural ... by adding the morpheme *-s*, is often seen as being only of peripheral importance. This evaluation is especially found

in older publications which categorize the *s*-plural mainly as being strange or "ignorant" and which advise people against using it. This is not always done as vigorously as in Heinsius (1818) who writes: "There is no *s*-plural ...," or Sütterlin (1918) who recommends that one should: "avoid the *s*-plural completely, or at least use it very carefully and at best only in foreign words." A detailed compilation about references of these and similar viewpoints can be found in Rettig (1972:99ff).

The situation is parallel in English, where the self-appointed language mavens and guardians of proper style have generally resisted the lawful regularization of denominal verbs, a robust grass-roots phenomenon (Kim, et al., 1991). This is not surprising: irregular words depend on memorization of forms witnessed in speech and writing, and hence will be most vulnerable in the least literate, and most in need of overt examples from style manuals and "experts" (e.g., familiar grammatical "hobgoblins" like *datum-data*, *criterion-criteria*, *alumnus-alumni*, and so on.) Regular affixation, because of its memory-independent default applicability, is bound to be seen by these nonpsychologically minded writers as the "lazy" way of doing inflection, and hence discouraged. Note that the facts of prescriptive injunctions regarding regular and irregular inflection run against the common connectionist assumption that the clear cases where people execute rules can be explained away as the conscious application of pedagogical injunctions (e.g., Smolensky, 1988).

But if children's unconscious language acquisition mechanisms naturally seek examples of rules (e.g., Pinker, 1984), there is ample evidence that an affix like *-s* is a default, despite its low frequency, for it could reveal its default status to the prepared mind in a number of ways (see Marcus, et al., 1992, for discussion). For example, when a child hears a pluralized name, the impossibility of storing it as a listed plural root tells the child that its affix must be acting as a default. Pluralized examples of onomatopoeia like *Kuckucks* "cuckoo" and *Wauwaus* "dogs", and of obvious noncanonical borrowings like

Café and *Kiosk*, can serve the same function. Furthermore, when a word has two plurals, one is usually *-s*, such as *Onkel-Onkel* "uncle-uncles" and *Onkel-Onkels*. In addition, an affix which applies in many different parts of morphophonological space is likely to be the default. Examples of the plural *-s*, unlike the other affixes, are haphazardly sprinkled throughout morphophonological space: to masculine, feminine, and neuter nouns, to words that are part of the canonical stress pattern and to those that aren't, to monosyllables and polysyllables, to both vowel-final and consonant-final stems. In contrast, the irregular affixes show systematic gaps in this table: *-er* never applies to feminines; zero only applies to nouns ending with a schwa syllable; *-en* applies mostly to feminines; other affixes do not appear with feminine polysyllables, and so on. We do not know which of these cues to defaulthood children depend on in learning affixes, but given one or two of these cues, the child could assign an affix default status and all of the phenomena in Tables 1 and 2 would fall out.

Are there broader psychological implications? The human mind is notable for its ability to override prototypes based on clusters of similar exemplars in memory. We recognize why 24683 is an odd number, and why Mick Jagger is a grandfather (Armstrong, Gleitman, and Gleitman, 1983), know that a smelly black offspring of racoons with a painted white stripe is not a skunk (Keil, 1989), apply *modus ponens* to brand new domains of knowledge (Smith, Langston, and Nisbett, 1992), joke that one cannot be a little bit pregnant, deny a beer to a responsible young man the day before his 21st birthday, and free an obviously guilty suspect on a technicality (Pinker and Prince, 1989). These abilities would seem to depend on categorizations based on formal conditions that override stored associations, exactly what a mental rule accomplishes. It is possible that regular rules in language and formal rules in categorization are examples of a kind of basic mental symbol-processing operation that is found in many domains of cognition.

The German plural system, notorious for its irregularity from Mark Twain to the present, might seem to be an unlikely home for a mental symbol-processing operation. Nonetheless we have shown through a variety of linguistic and psychological measures that it contains a clear default operation that, becaues of its rarity, cannot be just a very strong pattern in associative memory. In this way, German is the exception that proves the rule.

References

Anderson, S. (1992) *Amorphous morphology.* New York: Cambridge University Press.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983) What some concepts might not be. *Cognition, 13*, 263-308.

Aronoff, M. (1976) *Word formation in generative grammar.* Cambridge, MA: MIT Press.

Aronoff, M. & Sridhar, S. N. (1983) Morphological levels in English and Kannada. In J. F. Richardson, M. Marks, & A. Chukerman (Eds.), *Papers from the parasession on the interplay of phonology, morphology, and syntax.* Chicago: Chicago Linguistics Society.

Bellugi, U., Bihrle, A., Jernigan, T., Trauner, D., & Doherty, S. (1991) Neuropsychological, neurological, and neuroanatomical profile of Williams Syndrome. *American Journal of Medical Genetics Supplement, 6*, 115-125.

Berko, J. (1958) The child's learning of English morphology. *Word, 14*, 150-177.

Bornschein, M. & M. Butt (1987) Zum Status des *s*-Plurals im gegenwärtigen Deutsch. In: W. Abraham & R. Arhammar (Eds.) *Linguistik in Deutschland.* Akten des 21. Linguistischen Kolloquiums. Tübingen: Narr. 135-154.

Brinkmann, U. (1993) Unpublished doctoral dissertation, Department of Linguistics, University of Heidelberg, Germany.

Bybee, J. L. (1985) *Morphology: A study of the relation between meaning and form.* Philadelphia: Benjamins.

Bybee, J. L. (1991) Natural morphology: The organization of paradigms and language acquisition. In: T. Huebner and C. Ferguson (Eds.) *Crosscurrents in Second Language Acquisition and Linguistic Theories.* Amsterdam: Benjamins. 67-92.

Bybee, J. L. & C. L. Moder (1983) Morphological classes as natural categories. *Language, 59*, 251-270.

Bybee, J. L. & D.I. Slobin (1982) Rules and schemas in the development and use of English past tense. *Language, 58*, 265-289.

Chomsky, N. (1957) *Syntactic structures.* The Hague: Mouton.

Chomsky, N. (1965) *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. & Halle, M. (1968) *The sound pattern of English.* New York: Harper and Row.

Clahsen, H. & Marcus, G. F. (1992) The acquisition of inflectional rules: Evidence from German noun plurals. Paper presented at the Seventeenth Annual Boston University Conference on Language Development, October 23-25.

Clahsen, H. & Rothweiler, M (1992) Inflectional rules in children's grammars: evidence from the development of participles in German. *Morphology Yearbook*, 1-34.

Clahsen, H., Rothweiler, M. Woest, A., & Marcus, G. F. (1993) Regular and irregular inflection in the acquisition of German noun plurals. *Cognition, 45*, 225-255.

Daugherty, K. & Seidenberg, M. (1992) Rules or connections? The past tense revisited. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. pp. 259-264. Hillsdale, NJ: Erlbaum.

de Saussure, F. (1960) *A course in General Linguistics.* London: Peter Owen.

Fodor, J. A. (1968) *Psychological explanation.* New York: Random House.

Fodor, J. A. & Pylyshyn, Z. (1988) Connectionism and cognitive architecture: a critical analysis. *Cognition, 28*, 3-71.

Gordon, P. (1985) Level-ordering in lexical development. *Cognition, 21*, 73-93.

Halle, M. & Mohanan, K. P. (1985) Segmental phonology of modern English. *Linguistic Inquiry, 16*, 57-116.

Hare, M. & Elman, J. A connectionist account of English inflectional morphology: Evidence from language change. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. pp. 265-270. Hillsdale, NJ: Erlbaum.

Heinsius, T. (1818) *Kleine theoretisch-praktische deutsche Sprachlehre für Schulen and Gymnasien*. Berlin. (cited in Bornschein & Butt, 1987).

Helbig, G. & Buscha, J. (1987) *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht.* Leipzig. VEB Verlag Enzyklopädie.

Jackendoff, R. S. (1975) Morphological and semantic regularities in the lexicon. *Language, 51*, 639-671.

Janda, R. D. (1990) Frequency, markedness, & morphological change: on predicting the spread of noun-plural *-s* in Modern High German and West Germanic. In *ESCOL '90*, 136-153.

Keil, F. C. (1989) *Concepts, kinds, and cognitive development.* Cambridge, MA: MIT Press.

Kim, J. J., Marcus, G. F., Pinker, S., Hollander, M. & Coppola, M. (1992) Sensitivity of children's inflection to morphological structure. Submitted.

Kim, J. J., Pinker, S., Prince, A. Prasada, S. (1991) Why no mere mortal has ever flown out to center field. *Cognitive Science, 15*, 173-218.

Kiparsky, P. (1982) Lexical phonology and morphology. In I. S. Yang (Ed.) *Linguistics in the morning calm*. Seoul: Hansin, pp. 3-91.

Köpcke, K.-M. (1988) Schemas in German plural formation. Lingua 74: 303-335.

Kuczaj, S. (1977) The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior, 16*, 589-600.

Lachter, J. & Bever, T. G. (1987) The relation between linguistic structure and associative theories of language learning -- A constructive critique of some connectionist learning models. *Cognition, 28*, 195-247.

Lakoff, G. (1987) Connectionist explanations in linguistics: Some thoughts on recent anti-connectionist papers. Unpublished electronic manuscript, Internet.

Lieber, R. (1980) *On the organization of the lexicon.* Unpublished doctoral dissertation, Department of Linguistics and Philosophy, MIT, Cambridge, MA.

MacWhinney, B. & Leinbach, J. (1991) Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 40*, 121-157.

Marchman, V. (in press) Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience.*

Marcus, G. F. (1993) Negative evidence in language acquisition. *Cognition, 46*, 53-85.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J. & F. Xu (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57* (4, Serial No. 228).

McCarthy, J. & Prince, A. (1990) Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory, 8*, 209- 283.

McCarthy, J. & Prince, A. (in press) *Prosodic morphology.* Cambridge, MA: MIT Press.

McClelland, J. L. (1988) Connectionist models and psychological evidence. *Journal of Memory and Language, 27*, 107-123.

Meier, H. (1964) *Deutsche Sprachstatistik.* Hamburg.

Mugdan, J. (1977) *Flexionsmorphologie und Psycholinguistik.* Tübingen: Narr.

Neeleman, A. & Schipper, J. (1992) Verbal prefixation in Dutch: evidence for conversion. *Morphology Yearbook*, 57-92.

Newell, A. (1980) Physical symbol systems. *Cognitive Science, 4*, 135-183.

Newell, A. & Simon, H. A. (1981) Computer science as empirical inquiry: Symbols and search. In J. Haugeland (Ed.), *Mind design: Philosophy, psychology, artificial intelligence.* Cambridge, MA: MIT Press/Bradford Books.

Olsen, S. (1990) Konversion als ein konkatenativer Wortbildungsprozeβ. *Linguistische Berichte*, 127, 185-216.

Park, T.-Z. (1978) Plurals in child speech. *Journal of Child Language, 5*, 237-250.

Perlmutter, D. (1988) The split morphology hypothesis: Evidence from Yiddish. In: Hammond, M. & Noonan, M. (Eds.) *Theoretical Morphology.* New York: Academic Press.

Pesetsky, D. (in press) *Zero syntax.* Cambridge, MA: MIT Press.

Pfeffer, J. (1964) *Grunddeutsch: Basic (Spoken) German word list: Grundstufe.* Englewood Cliffs, NJ: Prentice Hall.

Pinker, S. (1989) *Learnability and Cognition: The Acquisition of Argument Structure.* Cambridge, MA: MIT Press.

Pinker, S. (1991) Rules of language. *Science, 253*, 530-535.

Pinker, S., & Prince, A. (1988) On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition. *Cognition, 28*, 73-193.

Pinker, S. & Prince, A. (1989) The nature of human concept: evidence from an unusual source. Unpublished manuscript, Department of Brain and Cognitive Sciences, MIT.

Pinker, S. & Prince, A. (1992) Regular and irregular morphology and the psychological status of rules of grammar. In L. A. Sutton, C. Johnson, & R. Shields (Eds.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on the Grammar of Event Structure.* Berkeley, CA: Berkeley Linguistics Society.

Plunkett, K. & Marchman, V. (1990) From rote learning to system building. Technical Report 9020, Center for Research in Language, University of California, San Diego.

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.

Prasada, S. & Pinker, S. (1993) Similarity-based and rule-based generalizations in inflectional morphology. *Language and Cognitive Processses, 8.*

Prasada, S. Pinker, S, & Snyder, W. (1990) Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated. Paper presented to the Annual Meeting of the Psychonomic Society, New Orleans, November 23- 25.

Pyles, T. & Algeo, J. (1982) *The origins and development of the English language.* (Third edition) New York: Harcourt Brace Jovanovich.

Raymond, E. S. (1992) *The new hacker's dictionary.* Cambridge, MA: MIT Press.

Rettig, W. (1972) *Sprachsystem und Sprachnorm in der deutschen Substantivflexion.* Tübingen: Narr.

Rumelhart, D. & J. McClelland (1986) On learning the past tenses of English verbs. Implicit rules or parallel distributed processing? In J. McClelland, D. Rumelhart and the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition.* MIT Press: Cambridge, MA.

Ruoff, A. (1981) *Häufigkeitswörterbuch gesprochener Sprache.* Tübingen: Niemeyer.

Schaner-Wolles, C. (1988) Plural vs. Komparativerwerb im Deutschen. Von der Diskrepanz zwischen konzeptueller und morphologischer Entwicklung. In: H. Günther (Ed.) *Experimentelle Studien zur Flexionsmorphologie.* pp. 155-186. Buske: Hamburg.

Schöler, H. & W. Kany. (1989) Lernprozesse beim Erwerb von Flexionsmorphemen: Ein Vergleich sprachbehinderter mit sprachunauffälligen Kindern am Beispiel der

Pluralmarkierung. In: G. Kegel et al. (Eds.) *Sprachwissenschaft und Psycholinguistik 3*. pp. 123-175. Westdeutscher Verlag: Opladen.

Seidenberg, M. (1992) Connectionism without tears. In S. Davis (Ed.), *Connectionism: Theory and practice.* New York: Oxford University Press.

Seidenberg, M. & Bruck, M. (1990) Consistency effects in the generation of past tense morphology. Paper presented to the Annual Meeting of the Psychonomic Society, New Orleans, November 23-25.

Seiler, Hansjakob (1970) Laut und Sinn: Zur Struktur der deutschen Einsilbler. In H. Steger (Ed.), *Vorschläge für eine Strukturale Grammatik des Deutschen.* pp. 414-428. Darmstadt, Wissenschaftliche Buchgesellschaft.

Selkirk, E. O. (1982) *The syntax of words.* Cambridge, MA: MIT Press.

Senghas, A., Kim, J. J., Pinker, S., & Collins, C. (1991) Plurals-inside-compounds: Morphological constraints and their implications for acquisition. Paper presented at the Sixteenth Annual Boston University Conference on Language Development, October 18-20, 1991.

Smith, E. E., Langston, C. & Nisbett, R. (1992) The case for rules in reasoning. *Cognitive Science, 16*, 1-40.

Smolensky, P. (1988) The proper treatment of connectionism. *Behavioral and Brain Sciences, 11*, 1-74.

Spencer, A. (1990) *Morphological theory.* Cambridge, MA: Blackwell.

Sproat, R. (1992) *Morphology and computation.* Cambridge, MA: MIT Press.

Sütterlin, L. (1918) *Die deutsche Sprache der Gegenwart. (Ihre Laute, Wörter, Wortformen und Sätze). Ein Handbuch für Lehrer und Studierende.* Leipzig. (cited in Bornschein & Butt, 1987.)

Twain, M. (1880/1979) The awful German language. Reprinted in *The Unabridged Mark Twain*. Philadelphia: Running Press.

Ullman, M., Hickok, G. & Pinker, S. (1992) Irregular and regular morphology in an aphasic. Paper presented at the Theoretical and Experimental Neuropsychology/Neuropsycholgie Experimentale et Theoreticale (TENNET) Conference, University of Quebec at Montreal, May 20-22.

Ullman, M. & Pinker, S. (1993) The English past tense system: Native speaker judgments of regular and irregular forms and the effects of frequency and similarity. Unpublished manuscript, Department of Brain and Cognitive Sciences, MIT.

van Dam, J. (1940) *Handbuch der deutschen sprache. Zweiter Band: Wortlehre.* Groningen: J. B. Wolter's Uitgevers-Maatschappij N. V.

Veit, S. (1986) Das Verständnis von Plural- und Komparativformen bei (entwicklungs) dysgrammatischen Kindern im Vorschulalter. In G. Kegel, et al. (Eds.) *Sprachwissenschaft und Psycholinguistik.* pp. 217-286. Westdeutscher Verlag: Opladen.

Wahrig, G. (1986) *Deutsches Wörterbuch.* Munich: Mosaik Verlag.

Williams, E. (1981) On the notions "lexically related" and "head of a word". *Linguistic Inquiry, 12*, 245-274.

Wiese, R. (1986) Schwa and the structure of words in German. *Linguistics, 24*, 695-724.

Wiese, R. (1987) Phonologie und Morphologie des Umlauts im Deutschen. *Zeitschrift für Sprachwissenschaft, 6*, 227-248.

Wiese, R. (1988) *Silbische und lexikalische Phonologie. Studien zum Chinesischen und Deutschen.* Tübingen: Niemeyer.

Wiese, R. (1992) *The phonology of German.* Unpublished manuscript, Heinrich-Heine-Universität Düsseldorf.

Wunderlich, D. (1986) Zum Aufbau lexikalischer Repräsentationen. Unpublished manuscript. Seminar für Allgemeine Sprachwissenschaft. Universität Düsseldorf.

Wunderlich, D. (1992) A minimalist analysis of German verb morphology. Arbeiten des Sonderforschungsberereichs 282: Theorie des Lexikons. Heinrich-Heine-Universität: Düsseldorf.

Wurzel, W. U. (1990) The mechanism of inflection: lexicon representations, rules, and irregularities. In W. U. Dressler, et al. (Eds.), *Contemporary Morphology*. Berlin: de Gruyter.

Xu, F. & Pinker, S. (1992) Weird past tense forms. Paper presented at the Seventeenth annual Boston University Conference on Language Development, October 23-25.

Zwicky, A. M. (1975) Settling on an underlying form: The English inflectional endings. In D. Cohen & J. Wirth (Eds.), *Testing linguistic hypotheses*. pp. 129-185. New York: Wiley.