

Regular and Irregular Morphology and the Psychological Status of Rules of Grammar

Author(s): Steven Pinker and Alan Prince

Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Grammar of Event Structure (1991), pp. 230-251

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Regular and Irregular Morphology and the Psychological Status of Rules of Grammar

Steven Pinker
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology

Alan Prince
Program in Linguistics and Cognitive Science
Brandeis University

Supported by NIH Grant HD18381 and NSF Grant BNS 91-09766. We thank our collaborators Chris Collins, Marie Coppola, Greg Hickok, Michelle Hollander, John J. Kim, Gary Marcus, Sandeep Prasada, T. John Rosen, Annie Senghas, Michael Ullman, and Fei Xu.

Lucretius could not credit centaurs;
Such bicycle he deemed asynchronous.
—William Empson, "Invitation to Juno"

For many years grammatical theory has often been seen as a branch of cognitive psychology, aiming to characterize the mental representations underlying knowledge of language. Nonetheless the psychological status of rules and principles of grammar remains unclear. Does a rule of grammar literally correspond to a data structure or computational procedure implemented in neural hardware? Or is it a mere human-readable compressed summary of linguistic judgment data, an epiphenomenon of neurocomputational processes of a very different character? Many other questions in cognitive science would feel the impact of an answer to this one, ranging from innateness to cognitive architecture to the causes of language development in childhood.

In this paper we report the early stages of a project that focuses on one well-defined phenomenon of grammar and treats it as a "model organism." We are hoping that intensive multidisciplinary study of all relevant aspects of the phenomenon will lead to a depth of understanding that would not be otherwise possible, and might help towards the resolution of longstanding theoretical oppositions by pinpointing which aspect of language competing theories might be right or wrong about. Specifically, we are studying the psychology of inflectional morphology, beginning with an intensive investigation of English past tense and plural marking.

Inflection nicely encapsulates the central theoretical issues concerning the psychological status of grammar, because it involves two subprocesses that are closely matched in function but very different in operation. *Regular inflection* (e.g., *walk-walked*) is perfectly rule-governed, and thus looks like a paradigm case of a grammatical rule implemented in human brains. Irregular inflection (e.g., *sing-sang*) shows varying degrees of unpredictability and thus would seem to involve brute-force memory: the English-speaking child hears his parents say (e.g.) *sang*, and at some point memorizes it. Because the regular process seems to be the very essence of the symbol-manipulating, algorithmic approach to language underlying most theories of grammar, whereas the irregular process seems to involve a quite

different kind of memory-driven processing, the issue of regularity/productivity serves to bring opposing approaches to language into sharp focus.

Principal Issues and Existing Theories

Traditional grammars and introductory textbooks, recognizing the difference in predictability between regular and irregular subsystems, hint that inflection involves two psychological mechanisms: rules for regulars, rote for irregulars. Though intuitively appealing, the theory is well-known to be inadequate, for three reasons.

Similarity between the morphological base and the (irregular) marked form. Rote memory, as a list of slots, is designed for pairs with unrelated members, like *bel/is*. But for subregular pairs in general, cross-linguistically, most of the phonological content of the stem is preserved in the inflected form. English *swing-sung* is a typical example (see Pinker & Prince, 1988, for extended discussion).

Similarity within the set of base forms undergoing a subregular process. An irregular alternation applies, often, to a family of phonetically similar items, even though the similarity is not enough to give necessary and/or sufficient conditions for the alternation. Thus, in English, we find irregular verb classes such as *sing, ring, spring, drink, shrink, sink, stink, begin, swim; keep, creep, deal, feel, kneel, mean, dream; grow, blow, throw, know, draw, fly; hit, put, cut*, and so on. Again, a rote list cannot explain why this huge degree of redundancy is maintained (Bybee, 1988; Pinker & Prince, 1988).

Semiproductivity. Irregular pairs are not a completely unproductive fixed list, but their patterns can to some degree be extended to new forms, on the basis of similarity to existing forms. All children occasionally make errors such as *bring-brang* and *bite-bote* (see Bybee & Slobin, 1982; Pinker & Prince, 1988). Irregulars can be added to the language diachronically by analogy with existing forms (i.e., *caught, cost, flung, knelt, quit, slung, stuck, strung*; Jespersen, 1942), a process that is especially obvious when many dialects are compared (e.g., *bring-brung, drag-drug, climb-clumb, heat-het*; see Mencken, 1936). Adult speakers of the standard dialect frequently extend irregular patterns to nonce stems in experimental tests (e.g., *spling-splung*; Bybee & Moder, 1983; Kim, Pinker, Prince, & Prasada, 1991; Prasada and Pinker, 1991).

The deficits of the traditional rote-&-rule theory have been handled in two very different ways. Generative phonology (e.g., Chomsky and Halle, 1968; Halle and Mohanan, 1985) implicitly accepts the view of memory as providing a set of (unrelated) slots. All regularities, of whatever generality, are to be abstracted out, and presented in the form of rules. In the classical version, this results in a lexicon that is the repository of pure unpredictability, reduced to an absolute minimum of specification; this is thought to subserve the psychological function of reducing the burden on memory (Bromberger & Halle, 1989). In more recent versions (e.g., forms of "Lexical Phonology"), the output of each lexical level is held to be listed, resulting in several entries for each item; presumably, the listing of information which is rule-generated is assumed not to place a burden on memory. In all versions of the theory, both regular and irregular processes are analyzed as manifestations of formally-similar rules, some of which are therefore limited in their scope by stipulations of various kinds.

This theory gives a straightforward account of input-output similarity: if an irregular rule is stated to replace a specified segment, the rest of the stem comes through in the output untouched, by default, just as in the fully regular case.

However, no real account is forthcoming for similarity within the input set and the associated semiproductivity of the alternation pattern. If a rule is simply attached to a list of the words it applies to, the similarity among stems within the list is unaccounted for. But if a rule is attached to a pattern shared by the words, by means of a context term for the rule or by means of a redundancy rule, it fails because the similarity to be accounted for is typically one of *family resemblance*, rather than necessary and/or sufficient conditions: if an "i \rightarrow Λ " rule were to be linked to a pattern like CC_ *ng* so as to embrace *string*, *fling*, *cling*, and so on, it will incorrectly include *bring-brought* and *spring-sprang* and fail to include *stick-stuck* and *spin-spun* (Bybee and Slobin, 1982; Pinker & Prince, 1988).

Connectionist models also collapse the regular-irregular distinction, but do so from the opposite end of the productivity scale, as it were, working from the general view that all processes, from the most idiosyncratic to the most regular, are functions of memory — using a much enriched (and perhaps more psychologically plausible) conception of what memory is. Under this conception, memory is associative and superpositional: individual items are dissolved into sets of features, and similar items literally overlap in their physical representations, sharing representational real estate. Learning consists of establishing excitatory or inhibitory connections between the atomic features of the representations, and may also depend on internal "hidden" features that a connectionist device can essentially define for itself. Many architectures now exist for structuring connectionist or "neural" networks, and they are demonstrably capable of performing a variety of nontrivial tasks.

The most complete attempt at modeling a real linguistic system, however, remains that of Rumelhart and McClelland (1986), which aims to simulate the acquisition of the English past tense. In it, a simple associative memory, implemented as a neural network, learns to associate uninflected inputs with the corresponding past tense outputs. A stem is decomposed into phonological features, each corresponding to a neuron-like unit in the input layer of nodes. Likewise, there is a layer of units representing the output inflected form; every input unit is directly connected to every output unit by a weighted link. Stem/past pairs are supplied to the network by a "teacher" and the model stores correlations among pairs of features in the stem and past by changing the weight of the links between their units. Link weights from different words are superimposed, and the output for a given word is computed by summing the weighted inputs and comparing the sum to a threshold. In this way the model generalizes to new verbs according to its similarity to old ones in the training set and according to the strength of the association between the stem features and various output features. There is no explicit representation of words or high-level rules, and a single network handles the regulars & irregulars: *-oss* in the input is linked to *-ossed* in the output in the same way that *-ing* in the input is linked to *-ang* in the output.

The achievement of the Rumelhart-McClelland model is that, after extensive training on several hundred verbs and their past tense forms, it successfully reproduces all the regulars and irregulars in the training set, and — more important — productively extends both regular and irregular alternations to many new test words, in accord with patterns of family resemblance.

The Rumelhart & McClelland model turns out to be more a general indication of what can be done with connectionist networks than a detailed scientific model of linguistic knowledge, performance, or acquisition. Pinker & Prince (1988), Lachter & Bever (1988), and Sproat (1992) point out a variety of severe

deficiencies in its treatment of linguistic structure. Stem/past similarity is totally unaccounted for, since the associative device can learn arbitrary input/output mappings (including linguistically impossible ones like string reversal), so long as they are relatively uniform across inputs. Lacking a representation of lexical items separate from their phonological content, the model cannot represent homophones with different past tense forms such as *ring-rang* and *wring-wrung*. By treating regulars and irregulars in a single associative mechanism, the model does not account for the fact that regular past tense formation, unlike irregular past tense formation, appears to be insensitive to frequency and similarity but can apply to any kind of form at all; low frequency and unusual-sounding regular verbs (e.g., *rhumba'd* and *anastomosed*) are readily inflected by English speakers, but not the Rumelhart & McClelland model (Prasada and Pinker, 1991). Universal inflectional properties of headed vs. headless (exocentric) structures are quite beyond the grasp of the model, since it intrinsically lacks notions like "head," "lexical item," and "morphological derivative." Nevertheless, the success of the model in picking up on patterns of stem-stem similarity and in using these productively — the very area where generative grammatical theory is largely silent — is striking, and suggests that the model, despite its manifest linguistic shortcomings, embodies a real insight into the organization of memory.

A New Approach

The conclusion we draw is that generative theories are fundamentally correct in their characterization of productive rules and structures, but deficient in the way they characterize memory of less predictable material, which must be associative and dynamic, somewhat as connectionism portrays it. It is necessary, then, to develop a new theory — an extension of the traditional rule & rote approach — which explicitly acknowledges the roles of rules on the one hand and of associative memory on the other. From such a theory, it follows that regular and irregular inflection must be computed by two different systems. Regulars are computed by an implementation of a classic symbolic rule of grammar, which concatenates an affix with a variable that stands for the stem. Irregulars are memorized pairs of words, but the linkages between the pair members are stored in an associative memory structure with certain connectionist-like properties (cf. Bybee, 1988). Thus while *string* and *strung* are represented as separate, linked words, the mental representation of the pair overlaps in part with similar forms like *shrink* and *bring*, so that the learning of *shrunk* is rendered easier given a constant number of learning trials, and analogies like *brung* occur with nonzero probability as the result of noise or decay in the parts of the representation that code the identity of the lexical entry.

The theory is close in spirit to theories in the generative literature that distinguish true productive rules of morphology from morpholexical "redundancy rules," which capture varying degrees of systematicity in lexical memory without freely licensing productive extensions to new forms (e.g., Jackendoff, 1975; Aronoff, 1976; Lieber, 1980). However, it aims to sharpen our psychological understanding of the phenomena attributed to the redundancy rules, which, qua rules, have always been somewhat mysterious in function. If our approach is correct, the reason they don't act like bona fide rules is that they are not rules at all, but epiphenomena of the way structured lexical entries are partially superimposed in memory.

The new theory, even in its preliminary form, allows us to integrate a large number of phenomena. Since it categorically distinguishes regular from irregular,

the theory predicts that the two types should dissociate from virtually every point of view. First, with respect to the psychology of language use, we predict that irregulars (as memorized items) must be affected by properties of associative memory, such as frequency and similarity, whereas regulars need not be. Second, with respect to language structure, we predict that irregulars (as listed items) should be available for other word-formation processes, whereas regulars — being nonlexical and freely created — should not be. Third, with respect to neural representation in actual brains, we predict that since regular and irregular are subserved by different mechanisms, it should be possible to find one system impaired while the other is untouched. The predictions can be tested using methods ranging from reaction time experiments to structural analysis of languages to study of child language acquisition to investigation of brain-damage and genetic language-deficits. We have been finding that existing evidence from all these sources seems to be converging on the predicted dissociations. Equally importantly from the point of linguistic research, the approach could lead to improvements in the separate theories of rule-action and of linguistic association. Only when the associationist complexities are removed from the rule system does it become possible to formulate clear principles of rule-form. Only when the truly rule-governed is abstracted from memory does it become possible to see the kind of architecture that associative memory for lexical structure should be endowed with.

In the rest of the paper we summarize evidence for this new theory, together with the major open questions. Though a coherent picture is beginning to emerge (see, Pinker, 1991), the work is very much in progress, and the global view will surely change as more of the details are uncovered.

Effects of the Laws of Association: Evidence for a Regular-Irregular Dissociation

Frequency effects. The new theory predicts that irregular forms are stored, and all generalizations of irregular patterns are directly read off the stored forms (possibly governed by a prototype, implicit in the pattern of superimposition in memory). In contrast, a rule is available to generate regularly inflected forms on-line, and prior exposure to and storage of the inflected form is not necessary and affords no crucial advantage. Since memory traces get stronger with additional exposures, the theory predicts that irregular past tense formation should be highly sensitive to frequency, but regular past tense formation in general should not be.

There is abundant evidence for the first prediction, that irregular morphology is highly frequency-sensitive. Lower-frequency irregulars are more likely to be overregularized by children in errors like *breaked* (both in spontaneous speech, and in elicitation experiments; see Marcus, Pinker, Ullman, Hollander, Rosen, and Xu, in press; Bybee and Slobin, 1982), to be uttered incorrectly by adults when under time pressure (Bybee and Slobin, 1982), to drop out of the language diachronically (Bybee, 1985), to coexist with regular versions in "doublets" in the standard adult dialect (e.g., *dived/dove*) (Stemberger, 1989, Ullman & Pinker, 1990), and within doublets, to be the less preferred member (Ullman & Pinker, 1990). Presumably as a result of these psychological effects, the irregular vocabulary of English consists predominantly of high-frequency words; indeed, the 13 most frequent English verbs are irregular (Francis and Kucera, 1982). Even within the set of unique irregulars in modern English, frequency exerts a clear effect on linguistic judgments. The effect can be seen clearly in idioms and clichés, because they can contain a verb that is not unfamiliar itself but

that appears in the idiom exclusively in the present or infinitive form. Irregular verbs in such idioms can sound strange when put in the past tense: Compare *You'll excuse me if I forgo the pleasure of reading your paper before it's published* with *Last night I forwent the pleasure of reading student papers*, or *I don't know how she bears the guy* with *I don't know how she bore the guy*.

In a number of ways, regular inflection is free from such effects. Beginning with linguistic data, we find that regular verbs in nonpast idioms do not sound worse when put in the past: compare *She doesn't suffer fools gladly* with *None of them ever suffered fools gladly*. Similarly, some non-idiomatic regular verbs like *afford* and *cope* have common stems but very low-frequency past tense forms (Francis & Kucera, 1982), presumably because usually appear with *can't* and other negatives or modals, which require the stem form. But the uncommon *I don't know how he afforded it (coped)* does not sound worse than *He can't afford it (cope)*.

The contrast can also be demonstrated quantitatively. Ullman and Pinker (1991) had 32 subjects rate the naturalness of the past and stem forms of several hundred verbs, both irregular and regular, each presented in a sentence in counterbalanced random order. For regular verbs (more specifically, regular verbs that did not rhyme with any irregular and hence are uniformly untainted by attraction to clusters of irregular verbs), ratings of their past tense forms correlate significantly (.62) with ratings of their corresponding stems, presumably reflecting the fact that the pasts are rule-analyzed for the occasion and hence simply inherit the naturalness of the stem, and the past tense ratings did not correlate significantly with the frequency of the past form (-.14, after statistically removing linear effects of the stem itself in a partial correlation procedure), presumably because prior encounters did not leave permanent memory traces that would have made the frequent past tense forms more natural. In contrast, ratings of irregular past tense forms correlate less strongly with their stem ratings (.32) but significantly with past frequency (.35, partialing out stem rating).

The contrast is also visible in patterns of production of inflected forms in real time. Stemberger & MacWhinney (1988) report somewhat contradictory evidence, some of which they interpreted as showing that higher frequency regulars were less error-prone in elicited production than lower frequency regulars. Unfortunately they compared total frequency, not past tense frequency; this confound makes it difficult to know whether the effects they did find were due to global difficulty in reading or uttering lower frequency verbs, or to past tense frequency per se. Prasada, Pinker, & Snyder (1990) selected pairs of verbs that had identical nonpast frequencies (i.e., each pair consisted of a regular verb and an irregular verb that were equated on the sum of the frequency of the stem form, the frequency of the *-ing* form, and the frequency of the *-s* form), but that differed in their past tense frequency. The stem was flashed on a screen, and subjects had to utter the past tense form; a voice-triggered switch measured response time. In three separate experiments lower past-frequency irregulars took significantly longer to utter than higher past-frequency irregulars, but there was no such difference for regular verbs. Control experiments had subjects produce 3rd-person-singular forms (regular for all the verbs used), and simply read aloud all the stems and inflected forms; data from these control conditions showed that the effect of frequency on the production of irregular past forms is not an artifact of the selection of easier stems in the irregular low-frequency group. The interaction between regularity and frequency is not a floor effect of frequency, since the regulars were on average *lower* in frequency than the irregulars and spanned a larger frequency difference (the

opposite of what a frequency floor effect would require), and since frequency effects are generally logarithmic, so bigger differences would be expected for the regulars.

Finally, in recognition, Stanners et al. (1979) obtained evidence that only the stem form of regular verbs is stored as a distinct lexical entry, and that the inflected forms were analyzed on line into stem plus affix, whereas both stem and past forms of irregulars were stored. They used the repetition priming effect with a word-nonword discrimination task: when subjects must decide whether a letter string is a word, they respond more quickly if they saw that word in the preceding trials. Earlier presentation of a regularly inflected form produced the same amount of priming as prior presentation of the stem itself, suggesting that the inflected form activated a mental lexicon entry for the stem, not a lexical entry for the inflected form. In contrast, irregular stems were primed to a much lesser extent by an earlier presentation of their inflected forms than by an earlier presentation of the stem. Stanners et al. also found that the number of letters shared by items in a pair did not predict the amount of priming found for irregularly inflected or derivationally related items, suggesting that the greater priming for regulars is not an artifact of the greater amount of letterwise overlap. Similar effects have been found by Kempley and Morton (1982), using detection of auditory words in noise; though see also Fowler, et al. (1985); Sandeep Prasada has partially replicated this effect, using several improvements in methodology and materials.

Effects of similarity. Cross-linguistically, irregular items fall into clusters that participate in the same or similar morphological alternations and that display family resemblances among the invariant portion of the stem. Diachronically, of course, this is often connected with divergence through sound change from a truly lawful initial situation. But higher-level phenomena like paradigm leveling and attraction into irregular clusters require a cognitive, acquisition-based explanation. We explain the persistence of family resemblance clusters as the product of enhanced retention, and occasional extension by analogy, on the part of learners when faced with similar sets of forms.

Such analogizing was studied by Bybee and Moder (1983), whose subjects were more likely to extend an irregular pattern to a nonce stem the more similar it was to an existing English cluster: *spling* is inflected as *splung* (cf. *string*, *cling*) more often than *spiv* is inflected as *spuv*. Prasada and Pinker (1991) replicated the effect and did a similar manipulation with regular verbs, for which the new theory predicts that no significant generalization gradient should occur. They created 3 sets of nonce verbs that contained a vowel participating in an irregular alternation, and that differed in their similarity to existing regular English verbs: "prototypical" *plip* resembles *flip*, *slip*, *trip*, *clip*, *nip*, etc.; "peripheral" *ploamph* has an initial CCV and a final VCC that occur in no English verbs; *smaig* is of an intermediate degree of similarity. In three experiments using different modes of presentation and kinds or ratings, we confirmed that the rated naturalness of regular past tenses of peripheral stems were no worse than those of prototypical stems, controlling for naturalness differences among the stems themselves.

This interaction is clearly predicted by the rule/associative-memory theory, which posits that regulars are generally formed by free concatenation of an affix with a stem, producing a form that should inherit the naturalness of the stem without any modification contributed by the frequency of pairing of the affix with similar kinds of stems. It appears to contrast with the behavior of pure associative models such as that of Rumelhart & McClelland (1986), which did not generalize indiscriminately across regulars; for example, it failed to produce any output for

novel verbs such as *jump* and *pump*. Suspecting that dissimilarity from items in the training set was the cause, Prasada duplicated Rumelhart & McClelland's training conditions and tested their model's ability to generalize to our sample of nonce forms. As expected, the model generalized irregular patterns to verbs according to similarity, like our subjects and those of Bybee and Moder (1983); but, unlike our subjects, it also analogized regular inflection according to similarity, producing regularly inflected forms only for verbs similar to those trained. Sproat (1992) reports simulations with Dana Egedi demonstrating comparable generalization failures in a hidden-layer back-propagation extension of the RM model; this suggests that it is not the perceptron architecture of the original model, now considered old-fashioned by most connectionists, that is responsible for its shortcomings at generalizing regular morphology.

Open Questions about Associative Effects

Are regularly inflected forms ever stored in memory? Though the new theory predicts that productive regular inflection does not depend for its success on storage of previously encountered regular forms, it does not deny that some such forms can in fact be stored. Indeed it would be an ad hoc stipulation to rule it out: if human memory is capable of storing an arbitrary irregular form as the past tense counterpart of a given verb stem, there is nothing that could prevent such memory slots from being filled with regularly inflected forms as a special case. Rather, the prediction is only that generalization does not *depend* on prior storage, and thus that there is generally no *need* for the system to store them. A corollary of this position is that storage *is* needed whenever an attested regular form could not have been predicted by the speaker's rule system. There are several such circumstances.

First, children, before they have formulated the language-specific parts of a given rule, have no choice but to memorize regularly inflected forms if they are to use them at all. Marcus et al. (in press), in their analysis of data from Courtney Cazden (1968) on children's use of regular and irregular inflection in spontaneous speech transcripts, found that young children, before they show clear evidence of possession of the regular rule in the form of overregularizations like *breaked*, sporadically use correct regular past forms in past tense contexts (in most such contexts, they leave the verb unmarked). Reliable tensing of regular verbs in obligatory past tense contexts occurs later and is correlated over time with overregularizations, suggesting that a transition from pure reliance on memory to availability of a rule underlies both developments.

Second, many regular plurals have developed meanings that do not correspond to the composition of the stem meaning plus plurality (such as, perhaps, *drinks* as a collective term for alcoholic beverages, and pluralia tantum forms like *alms* and *pants*), and they must be stored.

Third, if a verb is a doublet with equally-strong synonymous regular and irregular alternants, such as *dived/dove* or *slit/slitted*, then the very existence of the regular is unpredictable, because ordinarily the presence of an irregular form blocks the regular rule (Aronoff, 1976; Kiparsky, 1983; Pinker, 1984.) In such circumstances survival of both forms in a speaker's dialect can only be explained in terms of storage of both, and other data collected by Ullman and Pinker (1990), discussed below, support this prediction.

Finally, Ullman and Pinker (1991) discovered another, more surprising circumstance in which regulars are stored. For regular verbs that rhyme with at least one irregular (e.g., *blink*, which rhymes with *drink*, *stink*, and so on), past tense

frequency has a small but measurable positive effect on subjects' ratings of the naturalness of the past tense form, even when the global naturalness of the stem is partialled out. It is as if forms like *blinked* are somewhat unpredictable because of the attraction by analogy to the cluster of similar irregulars, which by itself would tempt the speaker to expect *blank* or *blunk*. The effect is tiny — even very rare regular verbs, such as *steep*, are rated much higher in their regular past tense form than the version analogous to a rhyming irregular pattern — and indeed we know from historical data that the effect must be tiny, given that reinforcement of an irregular past form by families of similar rhymes is not enough to prevent the verb from sliding into regularity as its frequency declines (Bybee, 1985). What we have found is only that speakers have some small tendency to remember those regular past forms that flout the weak attraction of irregular clusters.

The Nature of Associative Lexical Memory. The new theory, like associationist theories, proposes that irregular generalizations are based on memory-driven analogies, but it is not clear what this means in mechanistic terms. An associative network mapping between stem features and past features, a la Rumelhart and McClelland, is implausible because stem-past similarity is unexplained; such a model could easily learn unnatural mappings such as changing all a's to b's, all b's to c's, and so on (Pinker & Prince, 1988). Furthermore the actual behavior of their model showed that it blended outputs from bits and pieces of material associated with various stem properties, resulting in odd chimeras like *mail-membled* and *tour-toueder* — a distinctly nonlinguistic process.

An alternative that is far closer to generative models would associately link stem features in the input with a small number of rules or morphemes (e.g., vowel-changes, restricted suffixes) in the output. This would account both for the precision of the irregular mappings between stem and past (most of the stem preserved, and a small number of mappings existing across the irregular stems; see Pinker & Prince, 1988), as well as the family-resemblance and generalization-by-similarity effects: a rule would be applied with nonzero probability to stems similar to the ones it should apply to. However, this model has other empirical failings. The psycholinguistic evidence suggests that irregular pasts are stored as lexical entries, not computed on line. Similarly, Bybee and Slobin (1982) point out that speech errors occurring when irregular pasts are elicited are virtually always existing but incorrect English words (e.g., *rise-raise*), never novel rule products (e.g., *rise-rewse*). This suggests that the system stores word-word associations, not word-rule associations. Finally, the linguistic structure of the irregular classes is consistent with this conjecture: all 180 irregular stems, and their corresponding past tense counterparts, are monosyllables, the minimal word-size of English (see McCarthy & Prince, 1986, 1990, to appear), and there is more uniformity among irregular past forms than among mappings between stems and their past forms. For example, *-ought* is a frequent irregular rhyme, but it is associated with a heterogeneous set of stems: *buy, bring, catch, fight, seek, teach, and think*.

It is not clear exactly what kind of associative memory fosters just the kinds of analogies that speakers are fond of. Possibly a network of word-word associations might give rise the right generalization structure if the design of the lexical representation is informed by modern linguistic theory and its implementation is informed by models of superpositional memory. Here we can only present a rough sketch.

Words might be represented in a hierarchical hardware representation that separates stems and affixes, and furthermore distinguishes foot- and syllable-internal structure, finally representing segmental and featural composition at the

lowest level of units. Furthermore each of the possible contents of each representation would be implemented once as a single hardware "type"; particular words would be represented in separate "token" units with pointers to the types it contains. Links between stems and pasts would be set up during learning between their representations at two levels: between the token representations of each pair member, and their type representations at the level of representation that is ordinarily accessed by morphology: syllables, onsets, rhymes, feet (specifically, the structures manipulated in reduplicative and templatic systems, as shown in the ongoing work of McCarthy & Prince and others.) Ordinary correct retrieval results from successful traversal of token-token links; this would exhaust the process for pairs like *go-went* but would be reinforced by type-type links for members of consistent and high-frequency families like *sing-sang*. On occasions where token-token links are noisy or inaccessible and retrieval fails, the type-type links would yield an output that has some probability of being correct, and some probability of being an analogical extension (e.g., *brang*). Because the representation of input and output are each highly structured, such extensions would nonetheless be precise and follow constrained patterns, e.g., preserving portions of the stem such as onsets while substituting the appropriate rhymes, and avoiding the chimeras and fuzzy approximations that we do not see among real irregulars but that pure feature-to-feature networks are prone to making.

A more distant goal would be to see if such an interaction between linguistic representation and superimpositional memory can be applied to the notoriously capricious productivity of much of English derivational morphology, especially Latinate affixation and allomorphy attributed to Level 1 in Lexical Phonology, and other "partially productive" morphological processes.

Interactions among Lexical and Morphological Processes: Evidence for a Regular-Irregular Distinction

There is a large linguistic literature on interactions among processes of lexical storage, derivational morphology, and inflectional morphology (e.g., Aronoff, 1976; Kiparsky, 1983; Williams, 1981; Anderson, 1984; Lieber, 1980; Selkirk, 1982). Such phenomena provide particularly compelling interdisciplinary evidence for qualitative, universal, possibly unlearned structural organization of one part of the human language system and therefore can test one of the most contentious issues in language study. A crucial prediction, in simplified form, is as follows: irregulars are stored as lexical entries, and can be the *input* to lexical rules of morphology. In contrast, regulars are computed from lexical entries, so they operate on the *output* of other morphological rules. Two sets of morphological phenomena bear on the prediction.

Exocentrism. Pinker and Prince (1988) point out a fatal flaw in Rumelhart & McClelland's claim that past tense formation can be accomplished without representations of morphological structure, by mapping phonological features to phonological features. The flaw is that in real speakers, any irregular mapping is squelched in favor of the regular when the verb is sensed to be derived from a noun or adjective: *high-sticked*/**stuck*, *grandstanded*/**grandstood*; see Kiparsky, 1983). A simple principle explains this and related phenomena: irregularity is a property of stored roots, not words (which can be derived from roots by rules). A word can take an irregular form only if it is headed by an irregular root; otherwise, the regular rule applies, acting as the default. **High-stuck* is ruled out because its head is a

noun, not a verb, and noun roots cannot be listed in the lexicon as having any past tense, let alone an irregular form.

Moreover, only if the root is in "head" position (Williams, 1981; Selkirk, 1982) can its irregularity be inherited by the word as a whole. In English the head is the rightmost category at every level of word structure, and we posit that *all* features of a category — its pointer to a semantic referent, its grammatical category, its morphological features like animacy and gender, and its irregularity — are copied from the head at a lower level to the head at the next level up. If any such feature has not been copied, it diagnoses the word has not having a head, or being *exocentric*; the prediction is that if a word is exocentric by any criterion of non-inheritance of a root feature, it may not inherit the irregularity of the root either. This explains the regularization of the baseball term *flied-out*, which has a verb root *fly* but is immediately based upon the noun *a fly ball*. Its full structure is [IV|[NI|[IV|fly]]]; the intervening Noun category, since it cannot have been copied from the base verb, indicates exocentrism which also blocks the copying of the irregularity of that verb. Similarly *low-lifes*/**lives* can be explained in terms of the word not denoting a kind of *life* and hence being ineligible for the endocentric structure that would allow the irregularity of *life* to be inherited by the whole verb (similar explanations apply to *walkmans*, *the Mickey-Mouses in the administration*, *The Toronto Maple Leafs*, and many others; Pinker and Prince, in preparation.)

Kim et al. (1991) showed that subjects — including people lacking the benefit of a college education — were sensitive to this constraint when rating regular and irregular past tense forms of novel verbs derived from nouns or from other verbs. For example, when presented with novel irregular-sounding verbs (for example, *to line-drive*), they strongly preferred the regular past tense form (*line-driven*) when it was exemplified in a context that made it clear that it was based on a noun ("to hit a line drive"). In contrast, in a control condition for unfamiliarity where the items were based on existing irregular verbs ("to drive along a line"), the usual irregular was preferred. Kim et al. also showed that the effect is not reducible to degree of semantic similarity, and that usages that appear to be counterexamples are explained by the availability of alternative analyses to speakers: when and only when a verb can be given a sensible endocentric analysis (e.g., *sublet*, which can be analyzed either as [ν sub[ν let]] or as [ν [ν sublet]]), competing regular and irregular forms are permissible. Kim, Marcus, Hollander, and Pinker (1991) have demonstrated that children's inflectional processing is sensitive to exocentrism as well. For example, preschool children prefer the form *ringed* to describe putting a ring on a finger, and *rang* to describe an act of striking a resonant object (including nonprototypical instances of ringing).

Compounding. Kiparsky (1983) pointed out that compounds generally can contain irregular plurals as their nonhead members, but not regular plurals: *mice-infested*/**rats-infested*; *teeth-marks*/**claws-marks*; *purple-people-eater*/**purple-persons-eater*. This follows if compounds are formed only from stored words, not from complex rule products, and if irregulars but not regulars are stored words. Gordon (1986) showed that preschoolers are sensitive to this constraint when asked to form compounds like *X-eater*. The phenomenon involves high stakes, as Gordon showed that compounds containing any kind of plural are of near-zero frequency in English. Children's exclusive use of irregulars in the absence of input evidence exemplifying such a constraint, he argued, suggests that their morphological systems are innately organized so that irregular patterns are distinguished from regular ones and only the former can feed rules like

compounding; this is an example of Chomsky's "poverty of the stimulus" argument. Clahsen, Rothweiler, Woest, and Marcus (1991) report similar results from German children.

Open Questions about Lexical-Morphological Interactions

One complication for the proposal that compounding distinguishes irregular from regular plural nonheads is the existence of classes apparent counterexamples to Kiparsky's generalization in English, such as *drinks cabinet*, *parks commissioner* and *grants administration*; Senghas, Kim, Collins, and Pinker (1991) have collected scores of examples from speech and writing. Selkirk (1982) suggests that some of the compounded plurals are noncompositional forms with unpredictable meanings, hence must be stored despite their purely morphological regularity (e.g., *drinks* for alcoholic beverages). Others, perhaps, are metonymic, as when *parks* refers to the department or portfolio of parks. Unfortunately, although our list contains many forms that submit to these explanation, it also contains forms whose regular plurals seem perfectly compositional, such as *publications catalogue*, *injuries report*, *enemies list*, and *chemical weapons attack*.

An alternative hypothesis is that Kiparsky's constraint is not purely morphological but semantic, which then interacts with morphology. Say that compounds may contain only singulars and, by extension, collective terms (those referring to an aggregation as a whole). Some regular plurals meet this definition for semantic reasons, but all irregulars meet it for morphological ones, because as forms stored by virtue of their unpredictable morphology they are paired directly with their referent of a set of objects, rather than composed out of the meaning of the single referent plus the operator of plurality. Crosslinguistically, the idea of "number" is often reflected in semantic notions like "collectivity" and "distributivity" rather than in a simple, syntactically-agreeing numerosity distinction (Mithun, 1988; Bybee, 1985). Therefore one might argue that irregular plurals in languages like English, since they lie formally outside the productive plural system, are liable to interpretation in terms of these universal derivational categories. Tiersma (1982) shows specifically that noun referents that have natural collective interpretations are correlated with morphological irregularity diachronically and crosslinguistically.

Senghas, et al. (1991) put this hypothesis to the test and obtained conclusive refutation of it. Sixteen adults judged the naturalness of novel compounds like *geese-feeder* and *ducks-feeder* differing in whether the initial noun was regular or irregular (matched for semantics and frequency), singular or plural, affected collectively or noncollectively, and in a root or synthetic compound. Compounds containing irregular plurals were rated as far more acceptable than those containing matched regular plurals. Moreover, the effect was independent of collectivity: compounds where the referents of the nonhead word were affected as a collection were better than compounds where the referents of the nonhead word were affected individually, but the magnitude of this effect was perfectly additive with the effect of the regular/irregular contrast. If the effect of irregularity were only to foster a collective reading, then when such a reading was forced by our context instructions, the effect would have disappeared or at least have been reduced. Thus there is a robust, unconfounded regular/irregular distinction in compounding, just as Gordon and Kiparsky assumed.

Senghas, et al., also determined that the irregularity advantage held in root compounds and synthetic compounds with equal force, dashing any hopes that one

kind of compound was fussy about regularity and the other, the source of the counterexamples, was not. Currently it appears that at least a relativized version of the Kiparsky effect, namely a preference for irregular forms inside compounds, is a genuine part of human psychology, but it does not act to rule out all regular plurals-inside-compounds. Conceivably there are multiple compounding processes in English, a stringent but natural-sounding process that only accepts stored lexical entries (including, as an automatic special case, irregular inflected forms, but not regular inflected forms), and a less natural and less choosy process that accepts as first member virtually anything (including phrases or parts thereof as in *seat-of-the-pants executive* and *American history teacher* under one reading; see Lieber, 1988, Halle and Mohanan, 1985) this "loop" would embrace regularly inflected words as a special case.

Another complication involves differences in the interaction between compounding and plural inflection in other languages. Gordon (1986), prompted by Melissa Bowerman, noted that in Dutch, most nouns take the *-en* plural ending (phonetically schwa), yet many compounds contain *-en* plurals. Dutch also has an *-s* plural suffix, which does *not* appear in compounds (with one or two restricted kinds of exceptions). Gordon proposed that the *-en* is not "regular" in the crucial sense, in that it is not entirely predictable and may occasion idiosyncratic phonology, whereas the *-s* is fully regular and is the default, productive suffix. However, our own informant study, carried out by Chris Collins using a carefully constructed ratings questionnaire (Senghas, et al.), shows this to be unlikely. Both *-en* and *-s* were productively used by a native speaker to pluralize names, nonce-forms, and syntactic collocations; the two affixes have separate domains of productivity, defined by phonological criteria and by vocabulary class (e.g., foreign vs. native), but within those domains they are both demonstrably productive.

The Dutch situation, and the similar one for German, poses an unsolved but tantalizing problem. One possibility, suggested for German by Clahsen, et al. (1991) is the *-en* suffix is applied productively in the same block of rules as compounding ("Level 2" in Lexical Phonology), allowing it to feed compounding; the *-s* suffix applied downstream in Level 3, as in English, too late to feed compounding. An alternative is that feeding relations among the lexicon, derivation, and inflection are not dictated by a small number of levels or strata, but, in the unmarked case only simple lexical roots (including irregulars) may feed processes like compounding and derivational affixation, and the child relaxes the restrictions on a rule-pair-by-rule-pair basis, in the form of morphological selection restrictions on particular rules, as multiply-derived forms are recorded from the input. Fabb's (1988) analysis of the extremely limited number of permissible feeding relations in English derivation gives initial plausibility to such a suggestion.

Regular-Irregular Dissociations in Special Populations

If regular and irregular patterns are computed by different kinds of neural systems, we should expect them to dissociate in individual people impaired in specific ways: certain kinds of populations should evince the ability to produce irregulars but not regulars, or vice-versa.

Children. Marcus et al. (in press), confirming Ervin (1964) and Cazden (1968), showed that children younger than about 2 1/2 display one such dissociation: they produce correct irregular pasts and plurals before making overregularization errors. We showed that this famous "U-shaped" developmental curve is not a statistical artifact, and that it cannot be explained, as Rumelhart &

McClelland conjectured, as a response to an increasing mixture of regular verbs in the child's input. The proportion of regular verb tokens in children's and parents' speech remains unchanged throughout childhood, because high frequency irregular verbs (*make, put, take, etc.*) dominate conversation at any age. The proportion of regular verb types in children's vocabulary necessarily increases because irregular verbs are a small fraction of English vocabulary, but this growth does not correlate with overregularization errors. The traditional, and we argued, correct explanation for the developmental sequence is that memorization of irregular verb forms from parental speech can take place as soon as words of any kind can be learned, but deployment of the rule system must await the abstraction of the English rule from a set of pairs of words accumulated across different parental sentences and correctly juxtaposed as nonpast and past versions of the same lexical entry. The resulting time lag defines the "U"-shaped sequence. Evidence confirming that acquisition of the regular tense-marking process is the rate-limiting step in the appearance of the child's first errors is the fact that the onset and rate of overregularization correlate with the reliable use of regular past tense marking of *regular* verbs in obligatory past tense contexts.

Aphasics. Marin, Saffran, and Schwartz (1976) showed that two stroke victims with symptoms commonly characterized as "agrammatic aphasia" displayed frequent errors in reading inflected forms: *smiled* was pronounced as *smile*, *wanted* as *wanting*. They noted that "the difficulty ... is not found with irregular forms. Both irregular plural nouns and verbs with irregular past tense forms were read several orders of magnitude better than their regular counterparts." A more controlled comparison involved regular plural nouns such as *misers, clues, buds* and phonologically-matched counterparts that are not compositional plurals of singular terms, such as *trousers, news, suds*; the former were read at 90% accuracy; the latter at 45%. This is predicted if agrammatic symptoms are caused by damage to a neural system that involves circuitry for implementing the regular rule, which is necessary for analyzing the regularly inflected stimulus forms, but left the lexicon relatively undamaged, including stored irregulars, which could thus be successfully matched against the irregular stimuli. Greg Hickok has replicating the effect with a carefully-studied aphasic patient, using a list of regular and irregular verbs that were equated for stem frequency, past frequency, and pronounceability; control words with similar phonological content are also included.

Specific Language Impairment. Gopnik (1990a, 1990b, 1991) documents a syndrome of hereditary Specific Language Impairment whose most lasting grammatical symptom is difficulty in controlling inflectional features. The patients make frequent number, tense, agreement, person, and case errors in their speech, and in experiments, they have trouble converting present tense sentences to past tense sentences and in inflecting nonce verbs compared to their non-afflicted sibling controls. They performed far better with irregular verbs, occasionally producing them in speech, writing, and elicitation experiments, presumably because of their high frequencies. Interestingly as the dysphasics grew older they appeared to learn regular inflected forms one by one on a memorized basis: verbs with higher past tense than stem frequencies were learned first, and then additional regulars were learned one by one, with no transfer from one regular verb to another, in response to explicit teacher training. As adults the family members sound almost normal, though as mentioned, they had trouble inflecting nonce forms like the plural of *zoop*; this suggests an extensive process of memorizing regulars. In other words their ability to apply inflectional rules seems impaired relative to their ability to

memorize words: irregular forms are acquired relatively easily, enjoying their advantage of high frequencies, regulars are memorized as if they were irregular.

An interesting possible implication of the SLI and aphasic syndromes is that their deficits do not appear extend to the pure morphotactic operation of concatenating stem and affix. The aphasics and SLI children do not eschew all regularly inflected forms; their errors consist of incorrect combinations (e.g., *musics* and *Montreal Forums* in one of Gopnik's subjects) or combinations that are incorrect in context. Clahsen (1990) reports similar phenomena in German SLI children. If this is a general phenomenon, it would suggest that systematic encoding of grammatical features in less-than-word-size morphemes can be dissociated from handling the phonological material expressing those morphemes, and that it is the encoding process that is the neurologically vulnerable computation.

Williams Syndrome. It would be desirable to demonstrate a double dissociation between regular and irregular morphology, which would require a patient group that made frequent overregularization errors with irregulars, while having full control of the regular rule. According to the theory, we predict such a pattern to occur in a population whose grammatical system was intact, but who had problems in retrieving lexical items, especially high-frequency lexical items, from memory. Children with Williams syndrome (a syndrome of retardation probably caused by a defective gene expressed in the CNS involving calcium metabolism), as documented by Bellugi, et al. (1990), display these prerequisites: they are often described as "hyperlinguistic" or as having "selective sparing of syntax," but they retrieve words in a bizarre fashion, failing to answer questions with the high-frequency words offered by most other children, retarded or normal. For example, when asked to name some animals, they respond not with *dog, cat, pig* but with *unicorn, tyrandon, yak, ibex*. It almost seems too good to be true that some of these children, according to Bellugi et al., produced "morphological errors" as one of their few mistakes in grammar, which Klima and Bellugi (personal communication) confirm consist largely of overregularizations (16% of irregular past tense forms), a surprisingly high rate for children of this age (late teens).

Regular-Irregular Interactions and Blocking

Any theory positing alternative mechanisms for computing a form must specify how they interact. Following Kiparsky (1983), Aronoff (1976), and others, we assume a *Blocking* principle: retrieval of an irregular (e.g., *came*) blocks application of the regular rule (which would otherwise have yielded *comed*). The theory raises three interesting issues.

Regular-Irregular Doublets. The existence of *doublets* such as *dreamed/dreamt* at first glance poses a dilemma, because a blocking mechanism should proscribe them. Blocking cannot simply be abandoned, because we need it to explain the ungrammaticality of *comed* and *feeled*. Nor can one posit that learners are conservative recorders of all and only the past tense forms they hear, for this would leave them unable to form past tenses for low-frequency and nonce forms like *anastomose* and *wug*. Pinker (1984) argued on learnability grounds that learners store the regular past tense version of verb if and only if they also hear an irregular competitor to it. As discussed, this proposal is a simple corollary more general principle that the unpredictable must be memorized: a regular form that coexists with an irregular version would have been unpredicted by the speaker, given the operation of Blocking, and must be memorized, just like all irregulars are. Indeed, Ullman and Pinker (1990) found that subjects ratings of the naturalness of

both the regular and irregular members of doublets correlate with the respective frequencies of the past tense forms in English. In contrast, recall that regular past tense forms that do not belong (and are not attracted) to doublets are rated independent of frequency. This shows that hearing the regular version of a doublet (but not regular forms in general) is important to its survival in speakers.

This in turn raises the question of how doublets arise in a dialect to begin with. An obvious explanation runs as follows. If irregulars are stored in a frequency-sensitive memory, then the lowest-frequency irregulars, stored in weak traces, may provide the speaker with a barely above-noise memory signal, leaving them in doubt as to whether to block the regular rule. (For many people this is palpable when they are faced with deciding between, e.g., *striven* and *strived*; they report not being sure whether they have heard *striven*.) This theory predicts that the goodness of the irregular member of a doublet should be negatively correlated with the rating of the regular member, and this is what Ullman and Pinker (1990) found ($r = -.76$).

Overregularization in children. Another apparent counterexample to the operation of blocking — if thought of as an inherent aspect of the human linguistic apparatus — is children's apparent nonchalance in regularizing irregulars, as in *breaked*. For irregular verbs that a child has never heard before, regularization is the only choice, but children regularize irregular verbs that they have previously used correctly for months. Indeed, children frequently alternate between irregular and regularized past tense forms of a given verb in a single conversation. For this reason overregularization has been a major puzzle in language acquisition research. If children are happy with free variation between regulars and irregulars, what makes them turn into adults, who block regularization when they know an irregular? (Research on parent-child interaction shows that it cannot be overt corrections or other forms of parental feedback, for parents do not reliably react to their children's grammatical errors; see Morgan and Travis, 1989; Marcus, 1991). But if children's linguistic systems are designed to block overregularization, why do they do it so often?

Marcus, et al. (in press) addressed these questions in a massive study of the 11,500 irregular past tense forms in the spontaneous speech of 84 children. Their major discovery comes as a shock to all those who think of children as exception-hating relentless regularizers: Overregularization errors are at all ages a small minority of irregular past tense forms (2.5%). This finding, combined with the fact that overregularization rates for different verbs correlate negatively with the verbs' past tense frequency in parental speech, exorcises most of the paradox. We proposed a simple explanation: children, like adults, mark tense using memory (for irregulars) and an affixation rule that can generate a regular past tense form for any verb. Retrieval of an irregular blocks the rule, but children's memory traces are not strong enough to guarantee perfect retrieval. When retrieval fails, nothing blocks the rule, and overregularization results. The cure for overregularization is living longer, thereby hearing irregular past tense forms more often and strengthening their memory traces. If so, the single phenomenon of incomplete or unreliable blocking by low-strength memory traces for irregular verbs provides a single explanation for overregularization errors, adults' admission of new doublets, and diachronic loss of an irregular (see Bybee, 1985).

Real-time psychological processes implementing Blocking. An interesting psychological issue is how blocking might be computed in real time during speech production. The simplest model is that the speaker scans the list of irregulars for the target stem, outputs an irregular if one is found, else applies the regular as the

default. This predicts that irregular pasts will be produced in less time than regulars — exactly opposite to the finding of Prasada, Pinker, and Snyder (1990). Assuming their finding of a speed advantage for regular past forms holds up in a planned replication where (unlike the original studies) regular verbs do not constitute a majority of the experimental items, it would suggest the following simple alternative model. A stem is matched against the memory in which irregulars are stored, and fed into the regular rule mechanism, in parallel. A one-way lateral inhibitory pathway connects the irregular memory system to the regular rule system. Assuming that memory matching involves a stochastic, feature-by-feature comparison, the process of matching a stem against the irregular array would yield a continuous signal indicating the probability of a complete match, becoming more accurate over time as more features of the input and memory forms are compared; this signal is what inhibits the regular rule. Since determining that a match merely exists requires less information than recovering the specific content of the stored irregular, the regular system can be turned off before it outputs an overregularization, allowing the irregular to be produced once its entire form is retrieved. When there is no irregular match, the rule process runs to completion uninhibited and quickly outputs the regular form. This predicts that people will take longer to produce regular verbs that resemble irregulars, since an early false match signal will temporarily inhibit the rule process. Seidenberg and Bruck (1990) report preliminary evidence for such a finding.

Toward a Theory of Universal Morphology

Ultimately an understanding of the psychology of morphology will require a theory of "universal morphology" that predicts possible and impossible human languages and that characterizes the child's language learning mechanism. The strongest prediction is that we should find the following association: certain inflectional patterns should be predictable in form, readily applied to nonce stems regardless of their global similarity to existing stems, exclusively applicable to exocentric words (derived from other grammatical categories, borrowed from other languages, or with referents distinct from that of their roots), and inadmissible inside lexical compounds and other derived words. These, we claim, involve the psychological mechanism "regular rule." A language could also contain inflectional patterns that relate pairs of items each meeting the phonological definition of "canonical word" in the language, whose forms cannot be predicted perfectly by the form of the stem, that occur either in high frequency items or within families of similar stems or both, that are extended to new forms only tentatively and in cases of high similarity to existing stems, that may appear inside compounds and other derived words, and that may not apply to exocentric words. These are "associatively-memorized irregular pairs."

It is an extremely strong prediction that in any language one should find that phenomena in either of these two clusters should be found exclusively in association with one another, never in association with a phenomenon from the other cluster. Notice that this is not a prediction of Lexical Phonology, which does not establish any principled relation between productivity and assignment to lexical level, though the state of affairs can be represented within it. The prediction doubtless will turn out to be too strong, but even partial confirmations across languages would offer new insights into the role of predictability, productivity, and statistical patterns in influencing grammar and linguistic performance. Preliminary evidence that fits into the framework may be found in aspects of Yiddish

(Perlmutter, 1988), German (Clahsen, et al., 1991; Lieber, 1980), Dutch (Senghas, et al, 1991), Arabic (McCarthy & Prince, 1990), and Tagalog and Warlpiri (Lieber, 1980).

An especially interesting tenet of the theory is that high type frequency (i.e., majority status among the vocabulary items in a category), though virtually a definition of "regular" in traditional grammar, is *not* among the criteria for regularity in its psychological sense; the fact that the majority of English nouns and verbs are regular is largely an accident. In German, for example, it appears that the default (phonologically and lexically unrestricted, *ceteris paribus*) plural is *-s*, while the most common plural (by far) is *-en*. In Arabic, the most common plural type is an internal modification called the "iambic plural" in McCarthy & Prince's (1990) detailed study of the phenomenon; yet the default, used for noncanonical forms, is suffixation. To date, connectionist modelers have assumed an equation between "default" and "most common," and connectionist models are in fact extremely sensitive to frequency with which data is presented to them (Plunkett and Marchman, 1991). We believe that cases like German and Arabic default structure disconfirm the naive statistical interpretation of the notion "default" and will turn out to motivate the rule/memory distinction fundamental to our approach: defaults are rules, and are therefore learned and implemented by different mechanisms from other patterns of limited applicability.

Indeed, it is conceivable that the majority status of regular verbs and the default nature of the regular rule may have been dissociated in earlier stages of English, with their current confounding a contingent historical product. The strong verbs in Modern English are fossils of Indo-European ablaut classes, presumably governed by rules whose lawfulness was obliterated by sound changes by the time of Proto-Germanic. The weak *-d* suffix was introduced in Proto-Germanic and appears to have been used there and in its immediate descendants for derived forms and borrowings (Pyles and Algeo, 1982). Two distinctive traits of Modern English are that it borrowed many of its verbs from French and Latin (perhaps 60-65%), and that it formed large numbers of its other verbs from nouns (perhaps another 20%) — both of which are default circumstances that call for the regular suffix. (These rough estimates, computed by Michelle Hollander, are based on a sample of 200 verbs drawn randomly from the approximately 4000 in Francis & Kucera, 1982.) Thus the preponderance and heterogeneity of regular verbs in English might be an epiphenomenon of the rule-like default nature of the suffixing process, rather than vice-versa. If so, it would be getting it backwards to simulate the mental process of regular inflection by exploiting the statistics of modern English, and the strategy should fail outright in modeling the regular processes in languages like Arabic and ancestors of Modern English where the most frequent inflections do not correspond to the the default inflection.

Finally, the psychological correlate of a theory of universal morphology is a learning theory for morphology, currently all but nonexistent (see MacWhinney, 1978, Pinker, 1984, and Pinker & Prince, 1988, for first attempts). The two kinds of theories may be related as follows. The universally diagnostic contingencies between patterns of output forms and kind of underlying morphological process, taken from the work on Universal Morphology, are a suitable candidate for the cues that the child uses in determining whether he should store a pair of items in memory or coin a rule based on them, and whether or not to permit patterns of each type to feed or be fed by other morphological relations; such proposals can then be submitted to additional developmental tests (see, e.g., Pinker, 1984). Currently a variety of learning procedures of varying degrees of plausibility can be envisioned.

For example, a priori one might suppose that the child memorizes pairs, and elevates their shared pattern to "rule" status once the pattern is seen to apply to a large number of types, to a diversity of types, to words belonging to every competing pattern, and/or to exocentric forms. Crosslinguistic work, and examination of the linguistic input available to children by the time they appear to use rules, are required to weed out the obviously incorrect possibilities from such a list.

Similarly, one can ask whether the child determines the feeding relations in the grammar (a) by having Kiparsky's (1983) three levels innately built in; (b) by having a more general capacity for constructing layered morphology, such as a default in which only stored words feed morphological rules (Aronoff, 1976) and rules making syntactically-relevant distinctions are placed after all others (Anderson, 1984), which is then articulated and ramified as specific input words produced by multiple rule application are analyzed; or (c) only a large set of selection restrictions among pairs of rules, posited individually in response to the relevant complex form in the input, are posited by the child. Again, further crosslinguistic and developmental data are needed to distinguish the possibilities.

Conclusion

What, then, does the multidisciplinary study of regular and irregular morphology have to teach us about the psychological status of grammatical rules and principles? If the current conclusions hold up, they portray the regular affixation rule as a linguistic process with some interesting properties. The process appears to be modular, not directly governed by any aspect of real-world knowledge, nonassociative (unaffected by frequency and similarity), sensitive to highly abstract formal distinctions (root versus derived, noun versus verb), developing on a schedule not timed by environmental input, organized by principles that could not have been learned, possibly with a distinct neural substrate and genetic basis. Exactly these properties have been claimed for grammar as a whole (e.g., Chomsky, 1981), based on data concerning complex interactions among syntactic principles. Intriguingly, the claim also finds support in the details of a phenomenon as simple as the mental process of deriving *walked* from *walk*.

References

- Anderson, S. R. (1984) Where's morphology? *Linguistic Inquiry* 13, 571-612.
- Aronoff, M. (1976) *Word formation in generative grammar*. Cambridge, MA: MIT Press.
- Bellugi, U., Bihrlle, A., Jernigan, T., Trauner, D., & Doherty, S. (1990) Neuropsychological, neurological, and neuroanatomical profile of Williams Syndrome. *American Journal of Medical Genetics Supplement* 6, 115-125.
- Bromberger, S. & Halle, M. (1989) Why phonology is different. *Linguistic Inquiry* 20, 51-70.
- Bybee, J. L. (1985) *Morphology: a study of the relation between meaning and form*. Philadelphia: Benjamins.
- Bybee, Joan L. (1988) Morphology as lexical organization. In M. Hammond and M. Noonan, eds., *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego: Academic Press.

- Bybee, J. L. & Moder, C. L. (1983) Morphological classes as natural categories. *Language* 59, 251-170.
- Bybee, J. L. & Slobin, D. I. (1982) Rules and schemes in the development and use of the English past tense. *Language* 58, 265-289.
- Cazden, C. B. (1968) The acquisition of noun and verb inflections. *Child Development* 39, 433-448.
- Chomsky, N. & Halle, M. (1968) *The Sound Pattern of English*. New York: Harper and Row.
- Chomsky, N. (1981) *Lectures on government and binding*. Dordrecht, Netherlands: Foris.
- Clahsen, H. (1989) The grammatical characterization of developmental dysphasia. *Linguistics* 27, 897-920.
- Clahsen, H., Rothweiler, M., Woest, A, and Marcus, G. F. (1991) Regular versus Irregular Inflection in the Acquisition of German Noun Plurals. Submitted.
- 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.
- Fabb, N. (1988) English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory* 6, 527-539.
- 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.
- Gopnik, M. (1990a) Dysphasia in an extended family. *Nature* 344, 715.
- Gopnik, M. (1990b) Feature blindness: A case study. *Language Acquisition*, 1, 139-164.
- Gopnik, M. and Crago, M. (in press) Genetic dysphasia. *Cognition*.
- Gordon, P. (1986) 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.
- Jackendoff, R. S. (1975) Morphological and semantic regularities in the lexicon. *Language* 51, 639-671.
- Jespersen, O. (1942) *A modern English grammar on historical principles*, VI. Reprinted 1961: London: George Allen & Unwin Ltd.
- Kempey, S.T., and Morton, J. (1982) The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology* 73, 441-454.
- Kim, J. J., Marcus, G., Hollander, M., & Pinker, S. (1991) Children's Inflection is sensitive to morphological structure. *Papers and Reports in Child Language Development*.
- 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.
- Lachter, J. & Bever, T. G. (1988) The relation between linguistic structure and associative theories of language learning — A constructive critique of some connectionist learning models. *Cognition* 28, 195-247.
- Lieber, R. (1980) On the organization of the lexicon. Doctoral dissertation, Department of Linguistics and Philosophy, MIT.

- Lieber, R. (1988) Phrasal compounds in English and the morphology-syntax interface. Papers from the Parasession on Agreement in Grammatical Theory. CLS 24, 202-222.
- MacWhinney, B. (1978) Processing a first language: the acquisition of morphophonology. Monographs of the Society for Research in Child Development 43.
- Marcus, G. (1991) Negative evidence in language acquisition. Unpublished ms.
- Marcus, G., Pinker, S., Ullman, M., Hollander, J., Rosen, T. J., & Xu, F. (in press) Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*.
- Marin, O., Saffran, E. M., & Schwartz, M. F. (1976) Dissociations of language in aphasia: Implications for normal function. In S. R. Harnad, H. S. Steklis, & J. Lancaster (eds.), *Origin and evolution of language and speech. Annals of the New York Academy of Sciences* 280, 868-884.
- McCarthy, J. and Prince, A. (1986) *Prosodic morphology*. Widely circulated manuscript, Brandeis University and U. Mass. Amherst.
- McCarthy, J. and Prince, A. (1990) Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8, 209-283.
- McCarthy, J. and Prince, A. (to appear) *Prosodic morphology*. Cambridge, MA: MIT Press.
- Mencken, H. (1936) *The American Language*. New York: Knopf.
- Mithun, M. (1988) Lexical categories and the evolution of number marking. In M. Hammond & M. Noonan, (eds.) *Theoretical Morphology: Approaches in Modern Linguistics*. New York: Academic Press.
- Morgan, J. L., and Travis, L. L. (1989) Limits on negative information on language learning. *Journal of Child Language* 16, 531-552.
- Perlmutter, D. (1988) The split morphology hypothesis: Evidence from Yiddish. *Theoretical Morphology*. New York: Academic Press.
- Pinker, S. (1984) *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1991) Rules of language. *Science* 253, 530-535.
- 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. (in preparation) Of mice and men.
- 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) Generalization of regular and irregular morphological patterns by people and by networks. Submitted.
- 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.
- Pyles, T. & Algeo, J. (1982) *The origins and development of the English language*. (Third edition) New York: Harcourt Brace Jovanovich.
- Rumelhart, D. E. & McClelland, J. L. (1986) On learning the past tenses of English verbs. In J. L. McClelland & D. E. Rumelhart (eds.), *Parallel distributed processing: Volume 2*. Cambridge, MA: MIT 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.
- Selkirk, E. O. (1982) *The syntax of words*. Cambridge, MA: MIT Press.
- Senghas, A., Kim, J. J., Pinker, S., & Collins, C.. 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.
- Sproat, R. (1992) *Morphology and computation*. Cambridge, MA: MIT Press.
- Stanners, R. F., Neiser, J. J., Herson, W., P., & Hall, R. (1979) Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* 18, 399-41.
- Stemberger, J. and MacWhinney, B. (1988) Are Inflected Forms Stored in the Lexicon?. In M. Hammond & M. Noonan, (eds.) *Theoretical Morphology: Approaches in Modern Linguistics*. 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.
- Tiersma, P. (1980) Local markedness. *Language* 58, 832-849.
- 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) Connectionism versus symbolic rules in language: A case study. Paper presented at the Spring Symposium Series of the American Association for Artificial Intelligence, Stanford, March 26-28.
- Williams, E. (1981) On the notions "lexically related" and "head of a word". *Linguistic Inquiry* 12, 245-274.