The Past-Tense Debate

The past and future of the past tense

Steven Pinker and Michael T. Ullman

What is the interaction between storage and computation in language processing? What is the psychological status of grammatical rules? What are the relative strengths of connectionist and symbolic models of cognition? How are the components of language implemented in the brain? The English past tense has served as an arena for debates on these issues. We defend the theory that irregular past-tense forms are stored in the lexicon, a division of declarative memory, whereas regular forms can be computed by a concatenation rule, which requires the procedural system. Irregulars have the psychological, linguistic and neuropsychological signatures of lexical memory, whereas regulars often have the signatures of grammatical processing. Furthermore, because regular inflection is rule-driven, speakers can apply it whenever memory fails.

Steven Pinker*
Dept of Brain and Cognitive Sciences, NE20-413, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*e-mail: steve@psyche.mit.edu

Michael Ullman
Dept of Neuroscience, Research Building EP-04, Georgetown University, 3900 Reservoir Rd, NW, Washington DC 20007, USA.
e-mail: michael@georgetown.edu

For fifteen years, the English past tense has been the subject of a debate on the nature of language processing. The debate began with the report of a connectionist model by Rumelhart and McClelland [1] and a critique by Pinker and Prince [2], and has since been the subject of many papers, conferences and simulation models [3–7] (see also McClelland and Patterson in this issue [8]).

The past tense is of theoretical interest because it embraces two strikingly different phenomena. Regular inflection, as in *walk-walked* and *play-played*, applies predictably to thousands of verbs and is productively generalized to neologisms such as *spam-spammed* and *mosh-moshed*, even by preschool children [9]. Irregular inflection, as in *come-came* and *feel-felt*, applies in unpredictable ways to some 180 verbs, and is seldom generalized; rather, the regular suffix is often overgeneralized by children to these irregular forms, as in *holded* and *breaked* [10,11]. A simple explanation is that irregular forms must be stored in memory, whereas regular forms can be generated by a rule that suffixes *-ed* to the stem [12,13]. Rumelhart and McClelland challenged that explanation with a pattern-associator model (RMM) that learned to associate phonological features of the stem with phonological features of the past-tense form. It thereby acquired several hundred regular and irregular forms and overgeneralized *-ed* to some of the irregulars.

The past tense has served as one of the main empirical phenomena used to contrast the strengths and weaknesses of connectionist and rule-based models of language and cognition [8]. More generally, because inflections like the past tense are simple, frequent, and prevalent across languages, and because the regular and irregular variants can be equated for complexity and meaning, they have served as a test case for issues such as the neurocognitive reality of rules and other symbol-manipulating operations and the interaction between storage and computation in cognitive processing [5–7].

In this article we defend the side of this debate that maintains that rules are indispensable for explaining the past tense, and by extension, language and cognitive processes [3–5,14]. We review what the theory does and doesn't claim, the relevant evidence, the connectionist challenges, and our hopes for the future of the debate.

The Words-and-Rules theory

The Words and Rules (WR) theory claims that the regular–irregular distinction is an epiphenomenon of the design of the human language faculty, in particular, the distinction between lexicon and grammar made in most traditional theories of language. The lexicon is a subdivision of memory containing (among other things) the thousands of arbitrary sound–meaning pairings that underlie the morphemes and simple words of a language. The grammar is a system of productive, combinatorial operations that assemble morphemes and simple words into complex words, phrases and sentences. Irregular forms are just words, acquired and stored like other words, but with a grammatical feature like 'past tense' incorporated into their lexical entries. Regular forms, by contrast, can be productively generated by a rule, just like phrases and sentences. A stored inflected form of a verb blocks the application of the rule to that verb (e.g. *brought* pre-empts *bringed*). Elsewhere (by default) the rule applies: it concatenates -ed with the symbol 'V', and thus can inflect any word categorized as a verb (see Fig. 1).

Irregular forms, then, do not require an 'exception module'. They arise because the two subsystems overlap in their expressive power: a given combination of features can be expressed by words or rules. Thus either a word (irregular) or a rule-product (regular) can satisfy the demand of a syntactic or semantic representation that a feature such as past tense be overtly expressed. Diachronically, an irregular is born when (for various reasons) learners memorize a complex word outright, rather than parsing it into a stem and an affix that codes the feature autonomously [3].

http://tics.trends.com 1364-6613/02/$ – see front matter PII: S1364-6613(02)01990-3

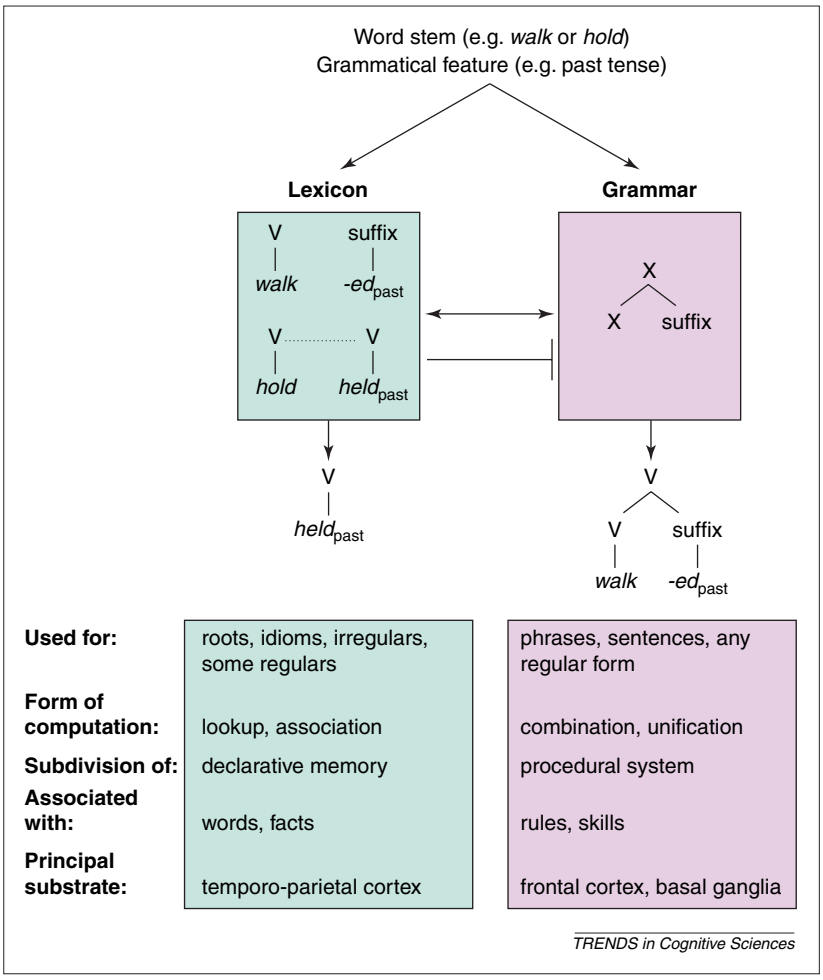

Fig. 1. Simplified illustration of the Words-and-Rules (WR) theory and the Declarative/Procedural (DP) hypothesis. When a word must be inflected, the lexicon and grammar are accessed in parallel. If an inflected form for a verb (V) exists in memory, as with irregulars (e.g. *held*), it will be retrieved; a signal indicating a match blocks the operation of the grammatical suffixation process via an inhibitory link from lexicon to grammar, preventing the generation of *holded*. If no inflected form is matched, the grammatical processor concatenates the appropriate suffix with the stem, generating a regular form.

The WR theory contrasts with classical theories of generative phonology and their descendents, such as those of Chomsky and Halle [15–17], which generate irregular forms by affixing an abstract morpheme to the stem and applying rules that alter the stem's phonological composition. Such theories are designed to account for the fact that most irregular forms are not completely arbitrary but fall into families displaying patterns, as in *ring-rang*, *sink-sank*, *sit-sat*, and *feel-felt*, *sleep-slept*, *bleed-bled*. A problem for this view is that irregular families admit numerous positive and negative counterexamples and borderline cases, so any set of rules will be complex and laden with exceptions, unless it posits implausibly abstract underlying representations (e.g. *rin* for *run*, which allows the verb to undergo the same rules as *sing-sang-sung*).

The theory also contrasts with the Rumelhart–McClelland model (RMM) and other connectionist models that posit a single pattern associator, with neither lexical entries nor a combinatorial apparatus [1,18,19]. The key to these pattern associators is that rather than linking a word to a word stored in memory, they link sounds to sounds. Because similar words share sounds, their representations are partly superimposed, and any association formed to one is automatically generalized to the others. This allows such models to acquire families of similar forms more easily than arbitrary sets, and to generalize the patterns to new similar words. Having been trained on *fling-flung* and *cling-clung*, they may generalize to *spling-splung* (as children and adults occasionally do [20,21]); and having been trained on *flip-flipped* and *clip-clipped*, they generalize to *plip plipped*.

WR is descended from a third approach: the lexicalist theories of Jackendoff, Lieber, and others, who recognized that many morphological phenomena are neither arbitrary lists nor fully systematic and productive [22–25]. They posited 'lexical redundancy rules', which do not freely generate new forms but merely capture patterns of redundancy in the lexicon, and allow sporadic generalization by analogy. Pinker and Prince proposed that lexical redundancy rules are not rules at all, but consequences of the superpositional nature of memory: similar items are easier to learn than arbitrary sets, and new items resembling old ones tend to inherit their properties. They argued that RMM's successes came from implementing this feature of memory, and proposed the WR theory as a lexicalist compromise between the generative and connectionist extremes. Irregulars are stored in a lexicon with the superpositional property of pattern associators; regulars can be generated or parsed by rules.

Ullman and colleagues have recently extended the WR theory to a hypothesis about the neurocognitive substrate of lexicon and grammar. According to the Declarative/Procedural (DP) hypothesis [5,26], lexical memory is a subdivision of declarative memory, which stores facts, events and arbitrary relations [27,28]. The consolidation of new declarative memories requires medial-temporal lobe structures, in particular the hippocampus. Long-term retention depends largely on neocortex, especially temporal and temporo-parietal regions; other structures are important for actively retrieving and searching for these memories. Grammatical processing, by contrast, depends on the procedural system, which underlies the learning and control of motor and cognitive skills, particularly those involving sequences [27,28]. It is subserved by the basal ganglia, and by the frontal cortex to which they project – in the case of language, particularly Broca's area and neighboring anterior cortical regions. Irregular forms must be stored in the lexical portion of declarative memory; regular past-tense forms can be computed in the grammatical portion of the procedural system.

What the words-and-rules theory does not say

The WR theory does not literally posit the discrete rule 'to form the past tense, add *-ed* to the verb'. All it posits is the past-tense morpheme *-ed*, a variable 'V' (included both in the attachment conditions for *-ed* and the lexical entry of every verb), and a general operation of merging or unifying constituents. The 'regular rule' or 'past-tense rule' is shorthand for the unification operation applied to the past-tense morpheme. WR is thus compatible with constraint- and construction-based theories of language, as long as they allow for variables and combinatorial operations [29].

WR does not posit that regular forms are *never* stored, only that they do not *have to* be [3,30–32]. It would be difficult to prohibit regular forms from ever being stored, given that human memory can acquire many kinds of verbal material (e.g. idioms, clichés, poems). WR posits a parallel-race model, like those defended for inflection by Baayen and Caramazza and by many psycholinguists for visual word recognition [33–39]. Whether a regular form is stored, and whether stored regular forms are accessed, depends on word-, task-, and speaker-specific factors [5,40–43]. For example, regular forms that constitute doublets with irregulars, such as *dived / dove* and *dreamed / dreamt*, must be stored to escape blocking by the irregular. As predicted, judgments of the naturalness of regular doublet forms show strong effects of frequency but other regular forms do not [30]. The same is true for regular forms of verbs that resemble irregulars (such as *blinked* and *glided*), because the forms must overcome a partial blocking effect exerted by the similar irregulars [30,32]. Tasks that require people to be sensitive to the physical form of words (such as progressive demasking) or to the prior existence of words (such as lexical decision), as opposed to tasks that ask people to judge possible forms, are likely to tap stored representations for medium- and high-frequency regular forms [3,35,44].

Finally, WR is not a chimera of a connectionist pattern associator glued onto a rule system. The lexicon has superpositional properties similar to a pattern associator, but lexical entries have structured semantic, morphological, phonological and syntactic representations of a kind not currently implemented in pattern associators.

Empirical tests

The key predictions of WR are: (1) that irregulars should have the psychological, linguistic and neuropsychological signatures of lexical memory, whereas regulars will often have the signatures of grammatical processing; and (2) that speakers should apply regular inflection whenever memory fails to supply a form for that category. A stored form may be unavailable for many reasons: low or zero frequency, lack of a similar form that could inspire an analogy, inaccessibility because of a word's exocentric structure (see below), novelty of the form in childhood, and various kinds of damage to the neurological substrate of lexical memory. The heterogeneity of these regular-eliciting circumstances offers converging evidence for distinguishable subsystems, including a productive default that does not critically depend on the statistics of patterns in memory. Here we discuss three types of evidence for a distinction between lookup and concatenation, and connectionists' attempts to provide alternative accounts (for reviews, see [3,4,14,31]).

Generalization to unusual novel words

The RMM model produced odd blends (*mail-membled*, *trilb-treelilt*), or no output, for novel words unlike those in its training set [2,20]. People, by contrast, readily apply regular inflections to novel unusual words [20]. According to WR, this is because *-ed* can attach to any word classified as a verb, even if dissimilar to existing stored regulars.

One connectionist explanation of the difficulties of the model is that they are specific to RMM, which is an early modeling exercise lacking a proper phonological representation, a hidden layer, and a proper output decoder. However, a pattern associator remedying all three deficiencies also had trouble generalizing to unusual words [45]. More recent models that are claimed to solve the problem do so, tellingly, by implementing or presupposing a rule. For example, Hare, Elman and Daugherty installed a 'clean-up network' in which the units for *-ed* strengthen the units for an unchanged stem vowel and inhibit the units for a changed vowel [46] – in effect, an innate mechanism dedicated to the English past tense. Many recent models have given up on generating past-tense forms; their output layer contains one unit for every past-tense change, turning inflection into a multiple-choice test among a few innate possibilities [47–49]. To convert the choice into an actual form, some other mechanism would have to copy the stem and apply the pattern corresponding to the selected unit. Such a mechanism is simply a rule. Marcus has argued that pattern associators' difficulty in generalizing to dissimilar forms is rooted in their design [4].

Another response is to claim that people's success at generalization depends on certain statistical patterns that also foster generalization in pattern associators. Many connectionists claimed that robust generalization depends on regular forms constituting the majority of forms in the child's input [50]. However, the onset and rate of over-regularization errors in children do not correlate with changes in the number or proportion of regular verbs used by parents [11,51,52]. Moreover, there are regular inflections in other languages, such as the German *-s* plural, that apply to a minority of nouns (~7%), but are generalized like English regular inflection, namely, to unusual nouns, exocentric nouns, and in childhood [50].

Several modelers now argue that it is not the number or proportion of regular words that is crucial but their distribution in phonological

Box 1. Systematic regularization

An intriguing aspect of inflection is that irregular forms can sometimes turn up in regular form. Some of these regularizations are unsystematic – for example, doublets such as *dived/dove* and *dreamt/dreamed*, in which the regular form is used sporadically because the irregular form is low in frequency and hence poorly remembered. But many are systematic: in particular contexts, the regular form is consistently used, such as *ringed the city* and *low-lifes*.

The Words-and-Rules theory explains this phenomenon using an independently motivated theory of compositionality in word-formation [a,b] (see also Fig. 2 in main article). Irregular-sounding words are regularized if they lack a root in head position that can be marked for the inflectional feature (tense or number). The regular suffix applies as the default, as it does in other cases where memory access is disabled. This neatly explains a diverse set of systematic regularizations found in actual usages, laboratory experiments with adults and children, and many languages [c–f]:

The word lacks a noun or a verb root
- **onomatopoeia:** *dinged*, *pinged*, *zinged*, *peeped*, *beeped*
- **quotations:** '*I found three* man's *on page 1*'; '*We to be'd and not to be'd in this room*'
- **names:** *the Julia Childs*, *the Thomas Manns*, *the Shelby Footes*
- **truncations:** *synched*, *sysmans*
- **unassimilated borrowings:** *talismans*, *mongooses*

The root cannot be marked for the feature
- **verbs with noun or adjective roots:** *ringed the city*, *steeled myself*, *spitted the pig*, *bared his soul*, *righted the boat*, *stringed the peas*
- **nouns with verb roots:** *a few loafs* (episodes of loafing), *a couple of wolfs* (wolfing down food)

The word's structure is exocentric
- **verbs based on nouns based on verbs:** *grandstanded*, *flied out*, *costed out the grant*, *encasted his leg*
- **nouns based on names based on nouns:** *Mickey Mouses* (simpletons), *Renault Elfs*, *Top Shelfs* (frozen food), *Seawolfs* (aircraft), *Toronto Maple Leafs*
- **nouns whose referents are distinct from those of their roots:** *low-lifes*, *still lifes*, *sabre-tooths*, *Walkmans*, *tenderfoots*
- **nouns based on phrases:** *Bag-A-Leafs*, *Shear-A-Sheeps*

Although the meaning of the regularized forms differs from that of their irregular counterparts, regularization is rarely triggered by differences in semantic features alone, as connectionists sometimes suggest [g,h]. If an irregular-sounding word changes in meaning, but retains a root in head position, it stays irregular, no matter how radical the change or opaque the metaphor:
- **compositional prefixing:** *overate*, *overshot*, *undid*, *preshrank*, *remade*, *outsold*
- **non-compositional prefixing:** *overcame*, *understood*, *withdrew*, *beheld*, *withstood*, *undertook*
- **compounding:** *bogeymen*, *superwomen*, *muskoxen*, *stepchildren*, *milkteeth*
- **metaphors:** *straw men*, *chessmen*, *snowmen*, *sawteeth*, *metrical feet*, *six feet tall*, *brainchildren*, *children of a lesser god*, *beewolves*, *wolves in sheep's clothing*
- **idioms:** *went out with* (dated), *went nuts* (demented), *went in for* (chose), *went off* (exploded), *went off* (spoiled); *took in* (swindled), *took off* (launched), *took in* (welcomed), *took over* (usurped), *took up*, (commenced), *took a leak* (urinated), *took a bath* (lost money), *took a bath* (bathed), *took a walk* (walked); *blew over* (ended), *blew away* (assassinated), *blew away* (impressed), *blew up* (exploded), *blew up* (inflated), *blew off* (dismissed), *blew in* (arrived)

[scores of other examples with *come*, *do*, *have*, *get*, *set*, *put*, *stand*, *throw*, etc.]

References

a Williams, E. (1981) On the notions of 'lexically related' and 'head of a word'. *Linguist. Inq.* 12, 245–274
b Selkirk, E.O. (1982) *The Syntax of Words*, MIT Press
c Pinker, S. (1999) *Words and Rules: The Ingredients of Language*, HarperCollins
d Kim, J.J. et al. (1994) Sensitivity of children's inflection to morphological structure. *J. Child Lang.* 21, 173–209
e Marcus, G.F. *et al.* (1995) German inflection: the exception that proves the rule. *Cogn. Psychol.* 29, 189–256
f Berent, I. *et al.* (1999) Default nominal inflection in Hebrew: evidence for mental variables. *Cognition* 72, 1–44
g Daugherty, K.G. *et al.* (1993) Why no mere mortal has ever flown out to center field but people often say they do. In *15th Annu. Conf. Cogn. Sci. Soc.*, Erlbaum
h Harris, C.L. (1992) Understanding English past-tense formation: the shared meaning hypothesis. In *Proc. 14th Annu. Conf. Cogn. Sci. Soc.*, Erlbaum

space [46,48,53,54]. If irregulars fall into clusters of similar forms (*sing*, *ring*, *spring*; *grow*, *throw*, *blow*; etc.), but regulars are sprinkled through no-man's-land, (*rhumba'd*, *oinked*, etc.), one can design pattern associators that devote some of their units and connections to the no-man's-land, and they will generalize to new unusual words. Putting aside the problem that most of these models have their inflections innately wired in, the models cannot deal with languages such as Hebrew, where regular and irregular nouns are intermingled in the same phonological neighborhoods. Nonetheless, Hebrew regular plural suffixes behave like *-s* in English and German: speakers apply them to unusual-sounding and exocentric nouns [55,56].

Systematic regularization

Some irregulars show up in regular form in certain contexts, such as *ringed the city* (not *rang*), *grandstanded* and *low-lifes* [2,57] (see Box 1 for further examples). This shows that sound alone cannot be the input to the inflection system: a given input, like *ring*, can be inflected either as *rang* or *ringed*, depending on some other factor.

The phenomenon falls out of the grammatical mechanism governing how complex words are formed [24,50,58,59]. Generally a complex English word inherits its features from its rightmost morpheme, its 'head'. For example, the head of *overeat* is *eat*; therefore, *overeat* is a verb (it inherits the 'V' category of *eat*), it refers to a kind of eating (because it inherits the semantic features of *eat*), and it has the irregular past-tense *overate* (because it inherits the stored past-tense form of *eat*) (see Fig. 2).

But there is a small family of exceptions: headless (exocentric) words, which for various reasons cannot get their features from their rightmost morpheme. For example, unlike endocentric verbs such as *overeat-overate* and *outdo-outdid*, which are verbs based on verbs, *to ring* and *to grandstand* are verbs based on nouns (*a ring*, *a grandstand*). In forming or parsing the word, the head-inheritance mechanism must be circumvented. With that data pathway plugged, there is no way for the irregular forms *rang*

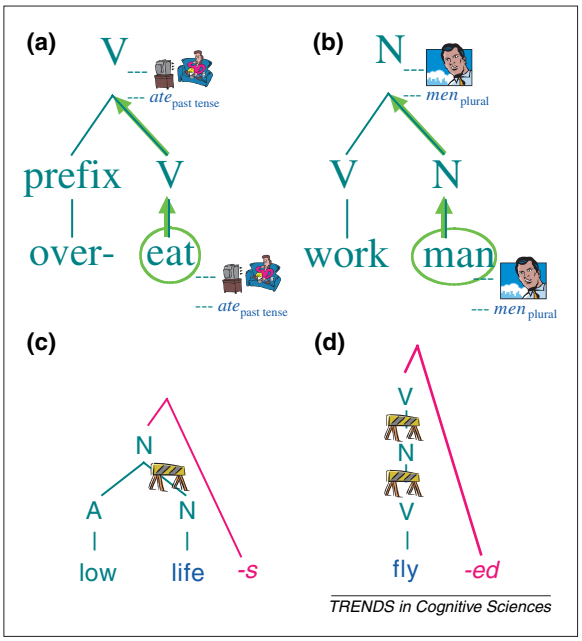

Fig. 2. Systematic regularization. Complex words are assembled out of simple morphemes according to a 'righthand-head' rule: the rightmost morpheme, the head, contributes its syntactic, semantic and morphological properties to the word as a whole. Thus in (a), the combination of *over-* and *eat* is a verb, because its head (circled), *eat*, is a verb (V); its meaning is a kind of eating (eating too much), because that is the meaning of *eat*, and its past-tense form is *overate*, because the irregular past-tense of *eat* is *ate*. All three kinds of information percolate up from the lexical entry for the head in memory along the rightmost edge of the word's tree structure (thick arrows). Similarly in (b), the combination of *work* and *man* is a noun (N), it refers to a kind of man, and its plural is *workmen*, the result of its inheriting all three properties from its head, *man*. However, a handful of derived words in English (headless or exocentric words) have to disable this inheritance mechanism. A *low-life* (c) is not a kind of life (in the way a *workman* is a kind of man) but a person who has a low life; for the word to work this way the usual data pipeline has to be blocked (depicted by the no entry sign). This leaves the irregular plural form (*lives*), trapped in memory, and the regular suffix *-s* applies as the default. The baseball term *to fly out* (d) comes from the noun *a fly* (as in *an infield fly*), which itself came from the simple verb root *to fly* (at the bottom of the tree). The word's structure requires the inheritance mechanism to be blocked twice: to allow the verb root *fly* to be converted to the noun (because verbs ordinarily beget verbs, not nouns) and again to allow the noun to be converted back into a verb (because nouns ordinarily beget nouns). The irregular past-tense forms *flew* and *flown* are sealed in memory, and *-ed* is suffixed as the default, generating *flied out*.

or *stood* to percolate up from the entries for *ring* or *stand*. With the irregular form sealed in memory, the suffixation rule steps in as the default, yielding *ringed* and *grandstanded*. Many examples, involving diverse constructions from several language families, have been documented from naturalistic sources and experimentally elicited from children and adults [3,50,60,61]. Apparent counterexamples exist, but virtually all can independently be shown to be cases where people do not assign an exocentric structure to the word [3,60].

There have been three connectionist explanations. One is that if a pattern associator had semantic as well as phonological input units, a complex word with an altered meaning would dilute the associations to irregular forms, favoring the competing regular [62,63]. However, in almost every case in which an irregular word's meaning changes, the irregular form is in fact *retained*, such as metaphors (*straw men*/**mans*, *sawteeth*, *God's children*) and idioms (*cut*/**cutted a deal*, *took a leak*, *hit the fan*, *put them down*) [2,3,50]. Accordingly, experiments have shown that just changing the meaning of an irregular verb does not cause people to switch to the regular [60,61]. Although all complex and derived words are semantically different from their bases, when semantic similarity and exocentric structure are unconfounded in a regression, exocentric structure accounts for a significant proportion of the variance in choice of inflectional form, and semantic similarity does not [60].

Equally unpromising is the suggestion that people regularize words to avoid ambiguity [63–65]. Many idioms are ambiguous between literal and idiomatic senses, such as *bought the farm* and *threw it up*, or among different idiomatic senses as well, such as *blew away* (impressed, assassinated), but this does not lead people to switch to a regular to disambiguate one of them (*buyed the farm*, *throwed up*). Conversely, *grandstood* and *low-lives* are unambiguous, but people still find them ungrammatical.

One connectionist model added nodes representing the semantic similarity of the verb to the homophonous noun (e.g. *to ring* and *a ring*) [64]. The network can then be trained to have these nodes turn off irregular patterns and turn on the regular one. But these unusual nodes are not part of the semantic representation of a verb itself; they are an explicit encoding of the verb's relation to the noun that heads it—that is, a crude implementation of morphological structure. In addition, the modelers had to train the network on regular past tenses of denominal verbs homophonous with irregulars. But such homophones are virtually absent from speech addressed to children, who nonetheless tend to regularize exocentric forms [61].

Neuropsychological dissociations

According to WR and DP, damage to the neural substrate for lexical memory should cause a greater impairment of irregular forms (and any regular forms that are dependent on memory storage), and a diminution of the tendency to analogize novel irregular-sounding forms according to stored patterns (as in *spling-splung*). In comparison, damage to the substrate for grammatical combination should cause a greater impairment of the use of the rule in regular forms, and of its generalization to novel forms.

Anomia is an impairment in word finding often associated with damage to left temporal/temporo-parietal regions (see Fig. 3a). Patients often produce fluent and largely grammatical speech, suggesting that the lexicon is more impaired than grammatical combination [66]. In elicited past-tense production tasks, patients (compared with controls) do worse with irregular than with regular verbs (Fig. 3b), produce regularization errors like *swimmed* (which occur when no memorized form comes to mind and the rule applies

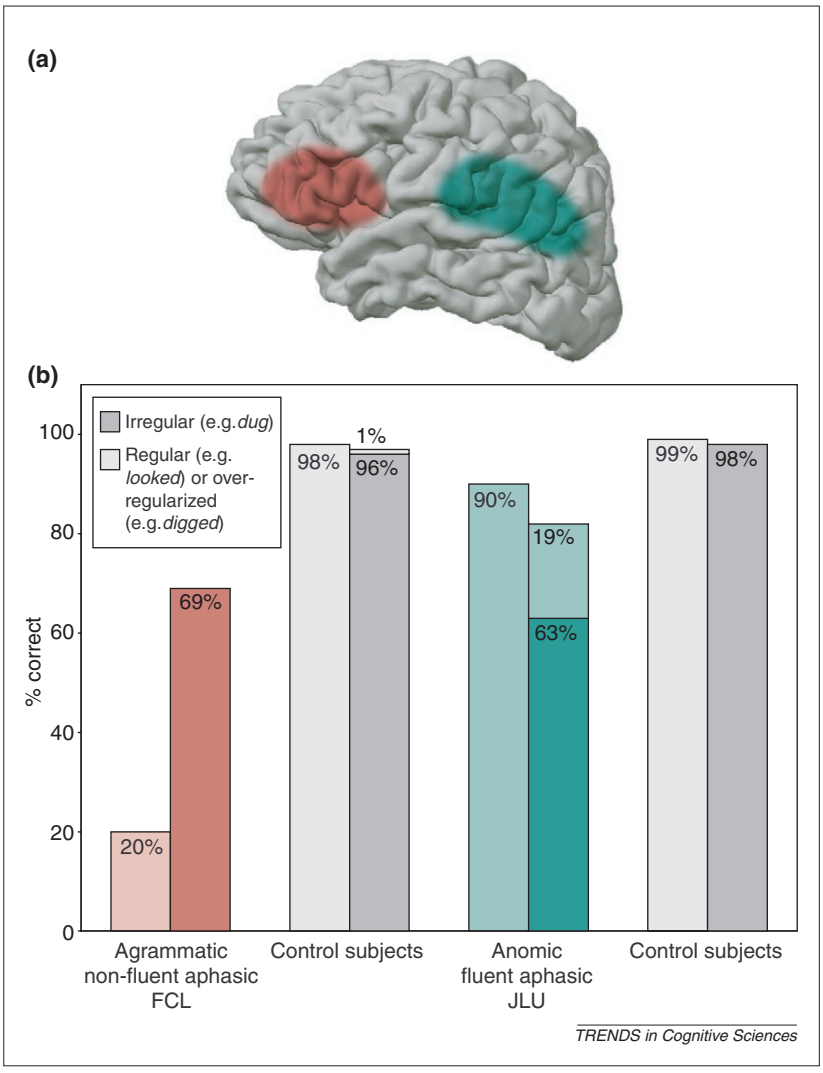

Fig. 3. Dissociating regular and irregular processing in aphasia. (a) The approximate lesion sites of patient FCL (red area, left anterior perisylvian regions), who had symptoms of agrammatism, and patient JLU (green area, left temporo-parietal region), who had symptoms of anomia. (b) Results of verb-inflection tests showed that the agrammatic patient had more trouble inflecting regular verbs (lighter bars) than irregular verbs (darker bars), whereas the anomic patient had more trouble inflecting irregular verbs – and overapplied the regular suffix to many of the irregulars (light green bar on top of dark green bar). The performance of age- and education-matched control subjects is shown in the grey bars.

as the default), rarely analogize irregular patterns to novel words (e.g. *spling-splung*), and are relatively unimpaired at generating novel regular forms like *plammed* [26,67,68]. Agrammatism, by contrast, is an impairment in producing fluent grammatical sequences, and is associated with damage to anterior perisylvian regions of the left hemisphere [69,70]. As predicted, agrammatic patients show the opposite pattern: more trouble inflecting regular than irregular verbs, a lack of errors like *swimmed*, and great difficulty suffixing novel words [26,67]. Similar effects have been documented in reading aloud, writing to dictation, repeating and judging words (even when controlling for frequency and length) [67],and in a regular/irregular contrast with Japanese-speaking patients [71].

The predicted double dissociation patterns are also seen in a comparison of neurodegenerative diseases. Alzheimer's disease (AD) is marked by greater degeneration of medial and neocortical temporal lobe structures than of frontal cortex (particularly Broca's area) and the basal ganglia, and greater impairment of lexical and conceptual knowledge than of motor and cognitive skills, including aspects of grammatical processing [72]. Parkinson's disease (PD), associated with basal ganglia degeneration, is marked by greater impairment of motor and cognitive skills (including grammatical processing) than use of words and facts [72,73]. As predicted, AD patients have more trouble inflecting irregular than regular verbs, are relatively unimpaired at suffixing novel words, generate few irregular analogies for novel words, and produce over-regularization errors; PD patients show the contrasting patterns [26,32]. Moreover, the performance patterns correlate with the severity of the associated processing impairments in the two populations: anomia in AD, and right-side hypokinesia (an index of left-hemisphere basal ganglia degeneration) in PD [26,32].

Intriguingly, Huntington's Disease (HD), caused by degeneration of different basal ganglia structures, results in disinhibition of the projected frontal areas, leading to unsupressible movements [73]. When HD patients inflect verbs, they show a third pattern: producing extra suffixes for regular and novel words like *walkeded*, *plaggeded* and *dugged*, but not analogous errors on irregulars like *dugug* or *keptet* – suggesting that these errors are instances of unsuppressed regular suffixation [26,32].

Converging findings come from other methodologies. In normal subjects, both regular and irregular inflected forms can prime their stems. By hypothesis, a regular form is parsed into affix and stem (which primes itself); an irregular form is associated with its stem, somewhat like semantic priming. Patients with left inferior frontal damage do not show regular priming (*walked-walk*), although they retain irregular priming (*found-find*) and semantic priming (*swan-goose*). A patient with temporal-lobe damage showed the opposite pattern [68,74,75]. In studies that have recorded event-related potentials (ERPs) to printed words, when a regular suffix is placed on an irregular word (e.g. the German *Muskels*) or omitted where it is obligatory (e.g. 'Yesterday I *walk*'), the electrophysiological response is similar to the Left Anterior Negativity (LAN) commonly seen with syntactic violations. When irregular inflection is illicitly applied (e.g. the German *Karusellen*) or omitted (e.g. 'Yesterday I *dig*'), the response is a central negativity similar to the N400 elicited by lexical anomalies, including pronounceable non-words [40,76–79]. This suggests that the brain processes regular forms like syntactic combinations and irregular forms like words.

Double dissociations are difficult to explain in pattern associators, because except for artificially small networks, 'lesioning' the networks hurts irregular forms more than regular ones [80]. A recent interesting model by Joanisse and Seidenberg

conceded that distinct subsystems have to be lesioned to produce double dissociations [81]. Although they called these modules 'phonological' and 'semantic,' the semantic module was in fact a lexicon: it had one unit dedicated to each word, with no representation of meaning. The finding that lesioning a lexicon differentially impairs irregular inflection is exactly what WR predicts. Moreover, the model failed to duplicate the finding that agrammatic patients have more trouble with regular than irregular verbs [26,67]. Lesioning the phonology module caused a consistent selective deficit only with novel verbs; regulars were no harder than irregulars. The report also claims that because a novel form has no meaning, 'the only way to generate its past tense is by analogy to known phonological forms' [81]. This predicts that patient groups should have parallel tendencies to generalize regular and irregular inflection to novel words (*plammed* and *splung*, respectively), whereas in fact these tendencies dissociate [32,67]. Finally, the model predicts that selective difficulty with irregular forms should depend on semantic deficits. Miozzo reports an anomic patient who had difficulty accessing word forms but not word meanings; nonetheless, he had trouble with irregulars but not with regulars [82].

The future of the past-tense debate

The Rumelhart–McClelland model was deservedly influential, we believe, because it captured a real phenomenon. The persistence of families of irregular verbs with overlapping partial similarities, and people's use and occasional generalization of these family patterns according to similarity and frequency, can be simply explained by the assumption that human memory is partly superpositional and associative. Theories that try to explain every instance of redundancy among words using the same combinatorial mechanism used for productive syntax and regular morphology require needless complexity and esoteric representations, and fail to capture the many linguistic, psychological and neuropsychological phenomena in which irregular forms behave like words.

At the same time, the post-RMM connectionist models have revealed the problems in trying to explain *all* linguistic phenomena with a single pattern-associator architecture. Each model has been tailored to account for one phenomenon explained by the WR theory; unlike RMM, few models account for more than one phenomenon or predict new ones. And modelers repeatedly build in or presuppose surrogates for the linguistic phenomena they claim to eschew, such as lexical items, morphological structure and concatenation operations. We predict that the need for structured representations and combinatorial operations would assert itself even more strongly if modelers included phenomena that are currently ignored in current simulations, such as syntax and its interaction with inflection, the massively productive combinatorial inflection of polysynthetic languages, and the psychological events concealed by providing the models with correct past-tense forms during training (i.e. children's ability to recognize an input as a past-tense form, retrieve its stem from memory, compute their own form, and compare the two).

As an increasing number of linguistic and neuropsychological phenomena are addressed, especially the complex data from neuroimaging, inadequacies will no doubt be revealed in both kinds of models. Nothing in linguistics prevents theories from appealing to richer conceptions of memory than simple rote storage. Neither does neural network modeling prohibit structured or abstract representations, combinatorial operations, and subsystems for different kinds of computation. The adversarial nature of scientific debate might sometimes have prevented both sides from acknowledging that features of one model may correspond to constructs of the other, described at a different level of analysis. We suspect that allowing a full range of data to tell us which processes are most naturally explained by which kinds of mechanisms, rather than shoehorning all phenomena into a single mechanism favored by one or another camp, holds the best hope for an eventual resolution of the past-tense debate.

Acknowledgements
We dedicate this paper to David Rumelhart and Jay McClelland, in acknowledgement of the stimulating effect of their model on psycholinguistic and neurolinguistic research. S.P. was supported by NIH grant HD-18381. M.T.U. was supported by a McDonnell-Pew grant in Cognitive Neuroscience, NSF SBR-9905273, NIH MH58189, and Army DAMD-17–93-V-3018/3019/3020 and DAMD-17–99–2-9007.

References

1 Rumelhart, D.E. and McClelland, J.L. (1986) On learning the past tenses of English verbs. In *Parallel Distributed Processing: Explorations in the Microstructures of Cognition* (Vol. 2) (McClelland, J.L. *et al.*, eds), pp. 216–271, MIT Press
2 Pinker, S. and Prince, A. (1988) On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193
3 Pinker, S. (1999) *Words and Rules: The Ingredients of Language*, HarperCollins
4 Marcus, G.F. (2001) *The Algebraic Mind: Integrating Connectionism and Cognitive Science*, MIT Press
5 Ullman, M.T. (2001) A neurocognitive perspective on language: the declarative/procedural model. *Nat. Rev. Neurosci.* 2, 717–726
6 Lima, S.D. *et al.* (eds) (1994) *The Reality of Linguistic Rules*, John Benjamins
7 Nooteboom, S. *et al.* (eds) (2002) *Storage and Computation in the Language Faculty*, Kluwer
8 McClelland, J.L. and Patterson, K. (2002) Rules or connections in past-tense inflections: what does the evidence rule out? *Trends Cog. Sci.* 6, 465–472
9 Berko, J. (1958) The child's learning of English morphology. *Word* 14, 150–177
10 Chamberlain, A.F. (1906) Preterite forms, etc., in the language of English-speaking children. *Mod. Lang. Notes* 21, 42–44
11 Marcus, G.F. *et al.* (1992) Overregularization in language acquisition. *Monogr. Soc. Res. Child Dev.* 57, 1–165
12 Chomsky, N. (1959) A review of B.F. Skinner's Verbal Behavior. *Language* 35, 26–58
13 Pinker, S. (1984) *Language Learnability and Language Development*, Harvard University Press
14 Clahsen, H. (1999) Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behav. Brain Sci.* 22, 991–1060
15 Chomsky, N. and Halle, M. (1968/1991) *The Sound Pattern of English*, MIT Press
16 Halle, M. and Mohanan, K.P. (1985) Segmental phonology of modern English. *Linguistic Inquiry* 16, 57–116
17 Halle, M. and Marantz, A. (1993) Distributed morphology and the pieces of inflection. In *The View from Building 20* (Hale, K. and Keyser, S.J., eds), MIT Press
18 Plunkett, K. and Marchman, V. (1993) From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition* 48, 21–69
19 Elman, J. *et al.* (1996) *Rethinking Innateness: A Connectionist Perspective on Development*, MIT Press

20 Prasada, S. and Pinker, S. (1993) Generalization of regular and irregular morphological patterns. *Lang. Cogn. Proc.* 8, 1–56
21 Xu, F. and Pinker, S. (1995) Weird past tense forms. *J. Child Lang.* 22, 531–556
22 Jackendoff, R.S. (1975) Morphological and semantic regularities in the lexicon. *Language* 51, 639–671
23 Aronoff, M. (1976) *Word Formation in Generative Grammar*, MIT Press
24 Lieber, R. (1980) On the organization of the lexicon. In *Linguistics and Philosophy*, MIT
25 Bresnan, J. (1978) A realistic transformational grammar. In *Linguistic Theory and Psychological Reality* (Halle, J.B. and Miller, G., eds), MIT Press
26 Ullman, M.T. *et al.* (1997) A neural dissociation within language: evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *J. Cogn. Neurosci.* 9, 266–276
27 Mishkin, M. *et al.* (1984) Memories and habits: Two neural systems. In *Neurobiology of Learning and Memory* (Lynch, G. *et al.*, eds), pp. 65–77, Guilford Press
28 Squire, L.R. and Knowlton, B.J. (2000) The medial temporal lobe, the hippocampus, and the memory systems of the brain. In *The New Cognitive Neurosciences* (Gazzaniga, M.S., ed.), pp. 765–780, MIT Press
29 Jackendoff, R. (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford University Press
30 Ullman, M.T. (1993) The computation of inflectional morphology. In *Department of Brain and Cognitive Sciences*, pp. 263, Massachusetts Institute of Technology
31 Ullman, M.T. (2001) The declarative/procedural model of lexicon and grammar. *J. Psycholinguist. Res.* 30, 37–69
32 Ullman, M.T. Evidence that lexical memory is part of the temporal lobe declarative memory, and that grammatical rules are processed by the frontal/basal-ganglia procedural system. *Brain Lang.* (in press)
33 Badecker, W. and Caramazza, A. (1991) Morphological composition in the lexical output system. *Cogn. Neuropsychol.* 8, 335–367
34 Burani, C. *et al.* (1984) Morphological structure and lexical access. *Vis. Lang.* 18, 348–358
35 Baayen, R.H. *et al.* (1997) Singulars and plurals in Dutch: evidence for a parallel dual-route model. *J. Mem. Lang.* 37, 94–117
36 Balota, D.A. and Ferraro, R. (1993) A dissociation of frequency and regularity effects in pronunciation performance across young adults, older adults, and individuals with senile dementia of the Alzheimer types. *J. Mem. Lang.* 32, 573–592
37 Balota, D.A. and Spieler, D.H. (1997) The utility of item-level analyses in model evalutation: a reply to Seidenberg and Plaut. *Psychol. Sci.* 9, 238–240
38 Besner, D. *et al.* (1990) On the connection between connectionism and data: are a few words necessary? *Psychol. Rev.* 97, 432–446
39 Coltheart, M. *et al.* (1993) Models of reading aloud: dual-route and parallel-distributed-processing approaches. *Psychol. Rev.* 100, 589–608
40 Ullman, M.T. *et al.* Sex differences in the neurocognition of language. *Brain Lang.* (in press)
41 Ullman, M.T. and Gopnik, M. (1999) Inflectional morphology in a family with inherited specific language impairment. *Appl. Psycholinguist.* 20, 51–117
42 van der Lely, H.K.J. and Ullman, M.T. (2001) Past tense morphology in specifically language impaired and normally developing children. *Lang. Cogn. Proc.* 16, 177–217
43 Ullman, M.T. (2001) The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingual. Lang. Cogn.* 4, 105–122
44 Alegre, M. and Gordon, P. (1999) Frequency effects and the representational status of regular inflections. *J. Mem. Lang.* 40, 41–61
45 Sproat, R. (1992) *Morphology and Computation*, MIT Press
46 Hare, M. *et al.* (1995) Default generalization in connectionist networks. *Lang. Cogn. Proc.* 10, 601–630
47 Hare, M. and Elman, J. (1992) A connectionist account of English inflectional morphology: evidence from language change. In *14th Annu. Conf. Cogn. Sci. Soc.*, pp. 265–270, Erlbaum
48 Nakisa, R.C. and Hahn, U. (1996) Where defaults don't help: the case of the German plural system. In *Proc. 18th Annu. Conf. Cogn. Sci. Soc.* (Cottrell, G.W., ed.), pp. 177–182, Erlbaum
49 Hahn, U. and Nakisa, R.C. (2000) German inflection: single route or dual route? *Cogn. Psychol.* 41, 313–360
50 Marcus, G.F. *et al.* (1995) German inflection: the exception that proves the rule. *Cogn. Psychol.* 29, 189–256
51 Marcus, G.F. (1995) The acquisition of inflection in children and multilayered connectionist networks. *Cognition* 56, 271–279
52 Marcus, G.F. (1995) Children's overregularization of English plurals: a quantitative analysis. *J. Child Lang.* 22, 447–459
53 Forrester, N. and Plunkett, K. (1994) Learning the Arabic plural: the case for minority default mappings in connectionist networks. In *16th Annu. Conf. Cogn. Sci. Soc.*, Erlbaum
54 Plunkett, K. and Nakisa, R. (1997) A connectionist model of the Arab plural system. *Lang. Cogn. Proc.* 12, 807–836
55 Berent, I. *et al.* The nature of regularity and irregularity: evidence from Hebrew nominal inflection. *J. Psycholinguist. Res.* (in press)
56 Berent, I. *et al.* (1999) Default nominal inflection in Hebrew: evidence for mental variables. *Cognition* 72, 1–44
57 Kiparsky, P. (1982) Lexical phonology and morphology. In *Linguistics in the Morning Calm* (Yang, I.S., ed.), pp. 3–91, Hansin, Seoul.
58 Williams, E. (1981) On the notions of 'lexically related' and 'head of a word'. *Linguist. Inq.* 12, 245–274
59 Selkirk, E.O. (1982) *The Syntax of Words*, MIT Press
60 Kim, J.J. *et al.* (1991) Why no mere mortal has ever flown out to center field. *Cogn. Sci.* 15, 173–218
61 Kim, J.J. *et al.* (1994) Sensitivity of children's inflection to morphological structure. *J. Child Lang.* 21, 173–209
62 MacWhinney, B. and Leinbach, J. (1991) Implementations are not conceptualizations: revising the verb learning model. *Cognition* 40, 121–157
63 Harris, C.L. (1992) Understanding English past-tense formation: the shared meaning hypothesis. In *Proc. 14th Annu. Conf. Cogn. Sci. Soc.*, Erlbaum
64 Daugherty, K.G. *et al.* (1993) Why no mere mortal has ever flown out to center field but people often say they do. In *15th Annu. Conf. Cogn. Sci. Soc.*, Erlbaum
65 Shirai, Y. (1997) Is regularization determined by semantics, or grammar, or both? Comments on Kim, Marcus, Pinker, Hollander & Coppola. *J. Child Lang.* 24, 495–501
66 Goodglass, H. (1993) *Understanding Aphasia*, Academic Press
67 Ullman, M.T. *et al.* Neural correlates of lexicon and grammar: evidence from the production, reading, and judgment of inflection in aphasia. *Brain Lang.* (in press)
68 Tyler, L.K. *et al.* (2002) Dissociations in processing past tense morphology: neuropathology and behavioral studies. *J. Cogn. Neurosci.* 14, 79–94
69 Alexander, M.P. (1997) Aphasia: clinical and anatomic aspects. In *Behavioral Neurology and Neuropsychology* (Feinberg, T.E. and Farah, M.J., eds), pp. 133–150, McGraw-Hill
70 Damasio, A.R. (1992) Aphasia. *N. Engl. J. Med.* 326, 531–539
71 Hagiwara, H. *et al.* (1999) Neurolinguistic evidence for rule-based nominal suffixation. *Language* 75, 739–763
72 Feinberg, T.E. and Farah, M.J. (eds) (1997) *Behavioral Neurology and Neuropsychology*, McGraw-Hill
73 Jankovic, J. and Tolosa, E.E. (1993) *Parkinson's Disease and Movement Disorders*, Williams & Wilkins
74 Marslen-Wilson, W.D. and Tyler, L.K. (1997) Dissociating types of mental computation. *Nature* 387, 592–594
75 Marslen-Wilson, W.D. and Tyler, L.K. (1998) Rules, representations, and the English past tense. *Trends Cogn. Sci.* 2, 428–435
76 Weyerts, H. *et al.* (1997) Brain potentials indicate differences between regular and irregular German plurals. *Neuroreport* 8, 957–962
77 Penke, M. *et al.* (1997) How the brain processes complex words: an event-related potential study of German verb inflections. *Cogn. Brain Res.* 6, 37–52
78 Gross, M. *et al.* (1998) Human brain potentials to violations in morphologically complex Italian words. *Neurosci. Lett.* 241, 83–86
79 Rodriguez-Fornells, A. *et al.* (2001) Event-related brain responses to morphological violations in Catalan. *Cogn. Brain Res.* 11, 47–58
80 Bullinaria, J.A. and Chater, N. (1995) Connectionist modeling: implications for cognitive neuropsychology. *Lang. Cogn. Proc.* 10, 227–264
81 Joanisse, M.F. and Seidenberg, M.S. (1999) Impairments in verb morphology after brain injury: a connectionist model. *Proc. Natl. Acad. Sci. U. S. A.* 96, 7592–7597
82 Miozzo, M. On the processing of regular and irregular forms of verbs and nouns: evidence from neuropsychology. *Cognition* (in press)

'Words *or* Rules' cannot exploit the regularity in exceptions

Reply to Pinker and Ullman

James L. McClelland
Center for the Neural Basis of Cognition and Dept of Psychology, Carnegie Mellon University, 115 Mellon Institute, 4400 Fifth Avenue, Pittsburgh, PA 15213, USA.
e-mail: jlm@cnbc.cmu.edu

Karalyn Patterson
Medical Research Council, Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, UK CB2 2EF.
e-mail: karalyn.patterson@mrc-cbu.cam.ac.uk

Pinker and Ullman [1] succinctly restate their position that the English past tense is governed by two competing mechanisms, identified as 'words and rules', and taken as examples of distinct procedural and declarative systems. Their mechanisms work separately, so that only one or the other is responsible for yielding a particular past tense form. To produce the past tense of *keep*, words and rules race to generate a response; as the correct past tense of *keep* is not regular, it must be generated by the lexical mechanism. For this reason, we think of their approach as the 'words *or* rules' theory.

Our approach is different. An integrated connectionist network maps from the stems of all verbs to their past-tense forms, using a single network of units and connections. For example, in the original Rumelhart and McClelland model [2], the same units and connections that produce regular past tenses from regular stems also process the irregulars, so the network has an inherent tendency to do the same thing to the exceptions that it does to regulars – namely, copy the features of the stem to the past-tense form and add /d/, /t/ or /^d/ depending on the final consonant. To produce *kept* instead of *keeped* (note that both end with unvoiced /t/) all that is required is to adjust the activations of the output units representing the vowel, something that the network will have learned to do on the basis of experience with *keep* and its neighbors *creep*, *leap*, *sleep*, *sweep* and *weep*. The network uses the same connection-based knowledge that allows it to perform the regular mapping, and also taps into specific connections activated by the particular properties of *keep* to produce the vowel adjustment.

A core difference between these approaches is that one exploits the regularity in the exceptions – what we call quasi-regularity – and the other does not. Quasi-regularity is the tendency for an exception to exhibit aspects of the regular pattern [3]. If there were only a few quasi-regular items, one might treat them as accidents, but in fact nearly all exceptional past-tenses in English are quasi-regular to some extent. To demonstrate this, we will review the different types (for other taxonomies, see [4,5]).

(1) Two very frequent verbs, *have* and *make*, delete a consonant and add the regular /d/ to what remains, forming *had* and *made*.
(2) The *-eep* words listed above and others, including *say*, *do*, *tell*, *sell*, *hear*, *flee* and *shoe*, form the past tense by adding regular /d/ or /t/ and making a vowel adjustment, producing *kept*, *said*, *did*, *told*, etc.
(3) Twenty-eight verbs, like *cut* and *hit*, have past tenses identical to their stems; all end in /d/ or /t/, as regular past tenses do.
(4) Another set of verbs ending in /d/ or /t/, including *bleed*, *breed*, *feed*, *lead*, *read*, *speed*, *hide*, *ride*, *slide* and *fight*, adjust the vowel to create /d/- or /t/-final *bled*, *slid*, *fought*, etc.

Several sets of verbs (waning in some dialects) use unvoiced /t/ instead of /d/, usually after /l/ or /n/:

(5) One such set, including *dwell*, *smell*, *spell*, *spill*, *burn* and *learn*, would be completely regular except for the de-voicing of the inflection, producing past forms like *spelt* and *burnt*.
(6) Another group, including *mean*, *dream*, *deal*, *feel* and *kneel*, adjust the vowel and add /t/, yielding *meant*, *dealt*, etc.
(7) A third set, including *build*, *bend*, *lend*, *rend*, *send* and *spend*, replace stem-final /d/ with /t/ to make *built*, *sent*, etc.
(8) Yet another set – *bring*, *catch*, *seek*, *teach* and *think* – adjust the vowel to /aw/ and replace the final consonant cluster with /t/, creating *brought*, *caught*, etc.

Overall, 59% of the 181 English exceptions listed by Pinker and Prince [5] have past tenses ending in /d/ or /t/, and fall into one of classes (1)–(8).

(9) Nearly all of the remaining verbs are also quasi-regular, in that the consonants of the stem are preserved. Instead of adding /d/ or /t/, the past tense is formed by making a vowel change, as in *sing-sang*, *rise-rose* and *fly-flew*.

There are only two 'suppletive' verb roots in English, *be* and *go*, with derivatives *forgo* and *undergo*, where the past-tense form is completely different from the present tense.

As noted above, the Pinker–Ullman theory provides no mechanism for exploiting the aspects of the regular past tense that are so prevalent among exceptions. Pinker did adopt the idea that the lexical system has connectionist-like properties [6]. This provided a way to account for clusters among the exceptions and for creative formation of novel forms consistent with such clusters. This was a step in the right direction, but did not go far enough. Because past tenses of exceptions in this account are formed by the lexical system alone, the theory still fails to explain why many of the exceptions share properties with regular past-tense forms and offers no way to exploit the regular mapping in forming past tenses of these exceptions.

By contrast, connectionist models inherently capture the regularity in the exceptions because the exceptions are processed by the same network

 1364-6613/02/$ – see front matter PII: S1364-6613(02)02012-0

that processes the regulars. As already noted for *keep-kept*, items that are quasi-regular can make partial use of the same connections that are used in forming exceptions. All nine of the types noted above, encompassing 177 out of 181 forms, exploit to some degree the connection weights that produce regular items. Only the suppletive items fail to make any use of the connections that produce the regular past tense [7].

The past tense of English is just one domain that exhibits quasi-regularity. In English spelling–sound mapping, virtually every exception has some degree of regularity; *pint*, *aisle*, *hymn* and *champagne* all partially adhere to regular correspondences. Quasi-regularity exists in richly inflected languages like Spanish, and in derivational as well as inflectional morphology [8,9]. It is found in language units beyond the word level [10,11] and, beyond language, it characterizes real-world objects, which have properties shared with other related objects as well as some unique properties [12]. Given these observations, the plausible candidate mechanisms of human linguistic and conceptual processes are those that can exploit quasi-regularity. Single-system connectionist models have this property; the Words *or* Rules theory does not.

References

1 Pinker, S. and Ullman, M. (2002) The past and future of the past tense. *Trends Cogn. Sci.* 6, 456–463
2 Rumelhart, D.E. and McClelland, J.L. (1986) On learning the past tenses of English verbs. In *Parallel Distributed Processing* (*Vol 2*): *Psychological and Biological Models* (McClelland J.L. and Rumelhart D.E., eds), pp. 216–271, MIT Press
3 Plaut, D.C. *et al.* (1996) Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol. Rev.* 103, 56–115
4 Bybee, J.L. and Slobin, D.L. (1982) Rules and schemas in the development and use of the English past pense. *Language* 58, 265–289
5 Pinker, S. and Prince, A. (1988) On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193
6 Pinker, S. (1991) Rules of language. *Science* 253, 530–535
7 Plunkett, K, and Marchman, V. A. (1991) U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition* 38, 43–102
8 Bybee, J.L. (1985) *Morphology: A Study of the Relation Between Meaning and Form*. John Benjamins
9 Burzio, L. (2002) Missing players: phonology and the past tense debate. *Lingua* 112, 157–199
10 Harris, C.L. (1994) Coarse coding and the lexicon. In *Continuity in Linguistic Semantics* (Fuchs, C. and Victorri, B., eds), John Benjamins
11 McClelland, J. L. (1992) Can connectionist models discover the structure of natural language? In *Minds, Brains and Computers* (Morelli, R. *et al.*, eds), pp. 168–189, Ablex Publishing
12 Rogers, T.T. and McClelland, J.L. *Semantic Cognition: A Parallel Distributed Processing Approach*. MIT Press (in press)

Rules or connections in past-tense inflections: what does the evidence rule out?

James L. McClelland and Karalyn Patterson

Pinker and colleagues propose two mechanisms – a rule system and a lexical memory – to form past tenses and other inflections. They predict that children's acquisition of the regular inflection is sudden; that the regular inflection applies uniformly regardless of phonological, semantic or other factors; and that the rule system is separably vulnerable to disruption. A connectionist account makes the opposite predictions. Pinker has taken existing evidence as support for his theory, but the review of the evidence presented here contradicts this assessment. Instead, it supports all three connectionist predictions: gradual acquisition of the past tense inflection; graded sensitivity to phonological and semantic content; and a single, integrated mechanism for regular and irregular forms, dependent jointly on phonology and semantics.

One view of language, originating with Chomsky [1,2], championed by Fodor and Pylyshyn [3] and widely pursued by Pinker [4–7], holds that abstract symbolic rules play a central role in human language processing. This claim is part of a broader view that human cognitive mechanisms are symbolic, modular, innate and domain-specific [4]. An alternative view, from Rumelhart and McClelland [8] (see Box 1), challenges the need for the use of rules. This view arises within the Parallel Distributed Processing (PDP) or connectionist framework [9], in which cognitive processes are seen as graded, probabilistic, interactive, context-sensitive and domain-general. Acquisition of language and other abilities occurs via gradual adjustment of the connections among simple processing units. Characterizations of performance as 'rule-governed' are viewed as approximate descriptions of patterns of language use; no actual rules operate in the processing of language.

These perspectives apply to many aspects of language, and, as Pinker and Ullman suggest [10], to many other domains as well, but here we focus on inflectional morphology, especially the English past tense. The idea of a past tense rule arose from noting that young children sometimes regularize irregular verbs, producing for example, *goed* or *felled* [11], and from the finding that children (and adults) typically produce regular forms for nonce (novel) words in a past-tense elicitation task [12]. Given a picture of a man said to be *ricking* and a request to complete '*Yesterday he* ___',

 1364-6613/02/$ – see front matter PII: S1364-6613(02)01993-9

Box 1. The Rumelhart–McClelland model

The Rumelhart–McClelland model of past-tense inflection [a] consists of a simple pattern-associator network [b,c] that learns the relationship between the phonological forms of the stems and past-tenses of English words. This network is flanked by a fixed encoding network on the input side and a fixed decoding network on the output side (see Fig. I). All learning occurs in the pattern associator. The encoding network simply converts a string of phonemes into the 'Wickelfeature' representation used inside the network to represent the stem of each word. Similarly, the decoding network converts the computed Wickelfeature representation of the attempted past-tense response back to a sequence of phonemes. The overall theory within which this model arose asserts that processing is meaning- and context-sensitive; for simplicity, such influences were not included in the model.

Processing

For a given input, the pattern associator produces an output by a simple neuron-like activation process. Each output unit computes a 'net input' based on the current input pattern and the values of the connection weights. The net input is the sum, over all of the incoming connections, of the activation of the sending unit multiplied by the weight of the connection. Each unit also has a modifiable threshold. When the net input exceeds the threshold, the unit tends to be turned on, with a probability approaching 1 as net input increases; otherwise, the unit tends to be turned off.

Learning

The network is trained using Rosenblatt's perception convergence procedure [d]. On a learning trial, the model is presented with the stem form of a word and its correct past tense. The stem form is encoded, and the activations of the Wickelfeature output units are computed. This computed representation is compared with the correct representation of the word's past tense. If the computed activation of a given unit matches the correct value, no learning occurs. If a unit that should be active is not, the weights to that unit from each active input unit receive a small fixed increment, and the threshold is reduced. Correspondingly, if a unit that should not be active is on, the weights from each active input unit are decremented and the threshold is increased. As a result, the network gradually improves performance over many learning trials, simulating a gradual developmental process. Later models use the back-propagation learning algorithm [e], an extension that allows the use of one or more layers of hidden units between inputs and outputs, and/or recurrent connections [f].

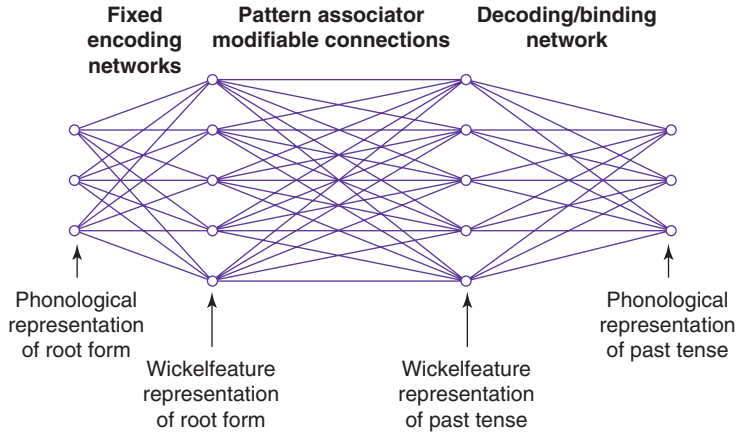

Fig. I. The Rumelhart–McClelland model of past-tense inflection (see text for discussion). Reprinted with permission from Ref. [a].

Representation

Coding is based on a idea by Wickelgren [g], in which word forms are represented by units designating each phoneme, together with its predecessor and its successor. Thus *help* would be represented by *_he, hel, elp*, and *lp_*. The model used units called 'Wickelfeatures' (WFs), each representing a feature from each of the phonemes in such triads. For example, there is a unit representing the feature sequence *liquid–unvoiced–end*, which would be active in representing *lp_*. In general, words ending in a unvoiced phoneme are represented by several WFs capturing the feature that the final phoneme is unvoiced. For the past tense output *helped*, such WFs should be replaced with others representing the added unvoiced stop /t/ that forms the past-tense inflection.

Capturing regular and exceptional inflections

For regular verbs in English, if the stem ends in a unvoiced sound (like the /p/ in *help*) the past tense will be formed by adding the unvoiced dental /t/. Through exposure to regular words, the network will repeatedly experience cases where the input contains WFs coding final unvoiced stem phonemes and the output contains WFs coding the added final /t/. The learning process will build up positive connections from the active input units to the appropriate output units, thereby encoding the regular addition of /t/ after unvoiced phonemes. Also, all non-final WFs of the stem are simply maintained in the past tense form, so the network will gradually acquire connections mapping each non-final WF to its counterpart in the output. At the same time, each output unit can be influenced by any input unit. To produce exceptions, connections from units coding specific input features to units coding for exceptional aspects of the inflection will be strengthened, thereby allowing specific properties of the input (such as presence of 'ee' followed by final /p/) to modify specific properties of the output, so that items like *creep, keep* and *sleep* are correctly mapped to the past tenses *crept, kept* and *slept*.

References

a Rumelhart, D.E. and McClelland, J.L. (1986) On learning past tenses of English verbs. In *Parallel Distributed Processing: Vol 2: Psychological and Biological Models* (Rumelhart D.E. and McClelland J.L ., eds), MIT Press

b Kohonen, T. (1977) *Associative Memory: A System Theoretical Approach*, Springer

c Anderson, J. (1977) Neural models with cognitive implications. In *Basic Processes in Reading: Percpetion and Comprehension* (LaBarge, D and Samuels, S.J., eds), pp. 27–90, Erlbaum

d Rosenblatt, F. (1962) *Principles of Neurodynamics*, Spartan

e Rumelhart, D.E. *et al*. (1986) Learning representations by back-propagating errors. *Nature* 323, 533–536

f Williams, R.J. and Zipser, D. (1989) A learning algorithm for continually running fully recurrent neural networks. *Neural Comput.* 1, 270–280

g Wickelgren, W.A. (1969) Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychol. Rev.* 86, 44–60

the response is usually *ricked*. As the child would never have heard *goed* or *ricked*, such responses were thought to show use of a rule.

We address a specific notion of rules held by Pinker and his collaborators, in which rules are discrete, categorical and symbolic objects used in a specialized, innate language module. For the English past tense, the rule takes as its argument any item identified only as a verb stem, and produces as its output its regular past tense. In English the output is stem + [d] (subsequent machinery realizes [d] as /d/, /t/ or /^d/, as in *loved*, *liked* or *hated*, depending only on the stem-final phoneme). The rule is said to be uniform in its application and

Table 1. Predicted and observed aspects of regular inflection

Aspect	Prediction from		Observed
	Symbolic Rules	Connectionist Models	
Acquisition	sudden	gradual	gradual
Sensitivity:			
to phonology	no	yes	yes
to semantics	no	yes	yes
in development	no	yes	yes
in German +*s* plural	no	yes	yes
Separability from exceptions:			
Genetically	yes	no	no
Neurologically	yes	no	no

James L. McClelland
Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University, 155 Mellon Institute, 4400 Fifth Avenue, Pittsburgh, PA 15213, USA.
e-mail: jlm@cnbc.cmu.edu

Karalyn Patterson
Medical Research Council, Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, UK CB2 2EF.
e-mail: karalyn.patterson@mrc-cbu.cam.ac.uk

independent of the meaning, phonology, frequency of occurrence, or any other attribute of the verb stem to which it applies. A further characteristic often attributed to such rules is that their acquisition is sudden. Thus Pinker suggests that the child 'deduces' the rule (Ref. [5], p. 193), calling this an 'epiphany' (p. 194) and a ' 'Eureka' moment' (p. 202). When we refer to symbolic rules, we mean rules with the characteristics just described.

Exceptions like *went*, *rang* and *slept* cannot be generated by the 'add [d]' rule. Pinker's theory proposes that they are dealt with by a lexical mechanism that is sensitive to frequency and similarity, and entirely distinct from symbolic rules. When planning to produce the past tense of a verb, the speaker first checks to see if an exceptional form can be retrieved from lexical memory. To account for the occasional occurrence of forms like *brang* (as the past tense of *bring*) or *splung* (as the past of the nonce verb *spling*), Pinker proposes that lexical memory has associative properties like PDP networks, and thus sometimes produces novel exception forms for inputs similar to known exceptions. In any case, if lexical memory offers up a form, it is produced; if not, the symbolic rule is used as a default. The theory encompassing the rule and the lexicon has been called the dual-mechanism or dual-route account.

Pinker and his colleagues, having examined several predictions of their account, conclude that the available evidence provides convincing support for it. The predictions are strong enough that confirmation would indeed support the idea of the symbolic rule mechanism. Furthermore, clear evidence for the purported properties of the symbolic rule mechanism would contradict basic tenets of the PDP alternative. The PDP account denies that rule-like aspects of language and other cognitive processes are generally characterized by the discreteness, uniformity of application, and modularity assumed for the symbolic rule system. It proposes that both regular and exceptional aspects of verb inflection (and of other aspects of language too; see [13,14]) emerge from a single, integrated mechanism. The connectionist approach makes opposite predictions to those of the rule-based approach (see Table 1), so that evidence against one is support for the other. It is therefore crucial to examine the evidence.

In what follows we consider whether inflectional morphology exhibits three key aspects of the symbolic rule (dual-mechanism) theory: (1) that acquisition of the symbolic rule is sudden; (2) that the rule is uniform in its applicability and independent of phonological, semantic or other factors; and (3) that the rule-based mechanism is separate from the mechanism that deals with exceptions.

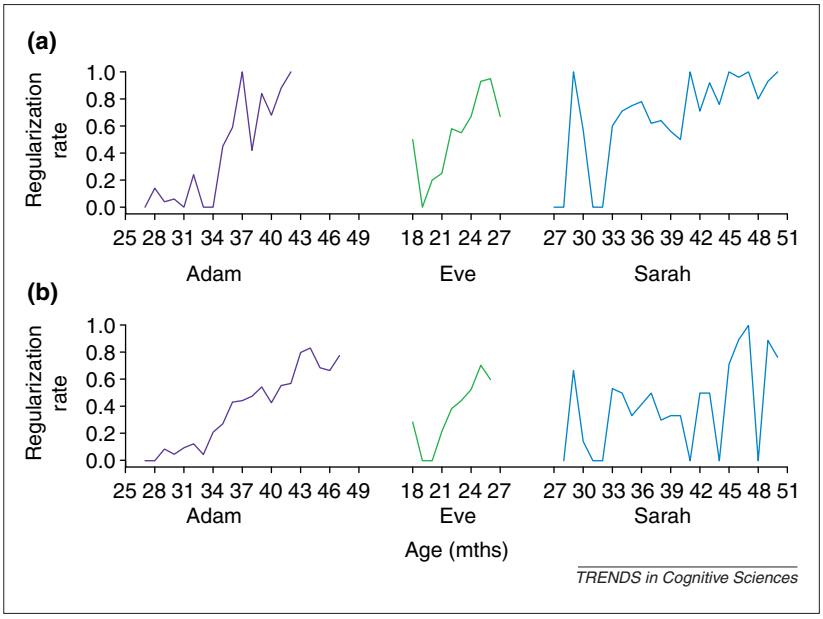

Fig. 1. Acquisition of the regular past tense by three children, Adam, Eve, and Sarah, as presented in Marcus *et al.* [15] and in Hoeffner (PhD thesis, Carnegie Mellon University, 1996). Percent usage of the regular past tense in obligatory contexts is plotted as a function of the child's age in months. (a) Marcus *et al.* presented data based on scoring by Cazden [16]. (b) Hoeffner repeated the analysis starting from the transcript provided in the CHILDES database [18], and included additional time periods. Two independent raters considered each occurrence of a regular verb in the child's speech, first considering the context of occurrence and evaluating whether a past tense was required before seeing the form of the verb actually used, thereby eliminating possible bias in determining whether the context required a past tense and producing an increase in the number of obligatory contexts identified. Data in (a) replotted based on data from Cazden [16] reprinted in Ref. [15].

Is acquisition of the regular past tense sudden?

Marcus *et al.* [15] considered the onset of the regular past tense, using Cazden's [16] analysis of recorded speech from three normally developing children (Adam, Eve and Sarah) [17]. Marcus *et al.* suggest that the first over-regularization in each child's corpus signals the moment of acquisition of the past-tense rule, and state that this over-regularization error is followed by 'rapid increases [in inflecting regulars] to high levels [...] shortly afterward. Adam's first over-regularization occurred during a 3-month period in which regular marking increased from 0 to 100%' (Ref. [15] , p. 103).

Hoeffner evaluated these data (J. Hoeffner, PhD thesis, Carnegie Mellon University, 1996), both as presented by Marcus *et al.* and as they emerged in a re-analysis using the transcription in the CHILDES database [18] (see Fig. 1). Considering first the data presented in Marcus *et al.*, Hoeffner noted that one could just as easily say that 'Adam's first

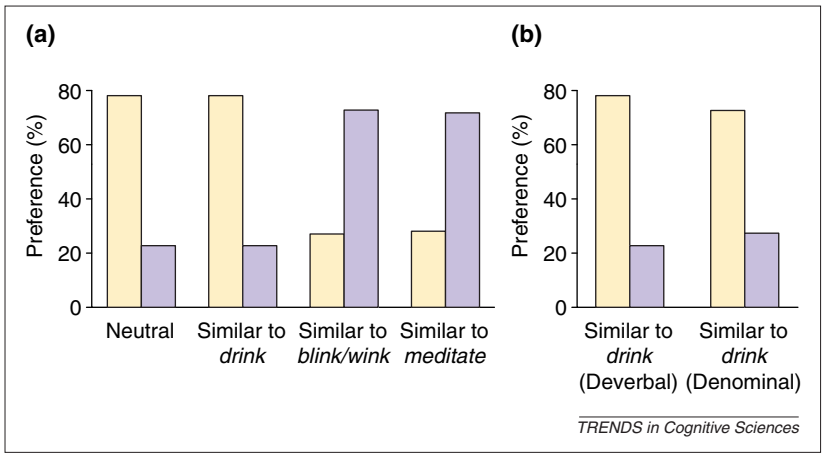

Fig. 2. Summary of effects of semantics and grammar on inflections of the nonce verbs *frink* and *sprink* from Ramscar [22]. (a) Use of irregular (*frank* or *sprank*, yellow bars) or regular (*frinked* or *sprinked*, mauve bars) in four different conditions. Note that in a neutral condition, with no semantic context, participants preferred irregular past tenses, and this trend persisted when context provided a meaning for the nonce verb similar to that of *drink*. When the context suggested a meaning similar to regular *wink* or *blink*, or even to the regular word *meditate*, participants shifted to the regular past tense, suggesting that use of the regular past tense can be influenced by semantics. (b) Subjects' ratings were not affected by their judgment of whether the nonce verb seemed to be denominal. Redrawn with permission from Ref. [22].

over-regularization occurred during a six-month period in which the probability of using the regular… rose gradually from 24 to 44%.' Either statement seems fairly arbitrary in fact; the data are noisy, and spikes occur when relatively few observations were available (Adam's 100% marking at 37 months is based on 8 observations). Given the noise, the graphs from all three children suggest a process that proceeds from very little marking in obligatory contexts to fairly reliable marking over the course of about one year. Hoeffner's own analysis (Fig. 1b), suggests an even more gradual acquisition process. A good fit to the data was achieved with a logistic regression in which the use of the regular past increases monotonically with age. Use of first over-regularization as a predictor did not reliably improve the account for regularization rates in any of the three children.

In short, the acquisition of the regular past tense is not sudden. According to Brown, reviewing Cazden's analysis of other inflections, the situation is the same in all cases:

> There is always a considerable period… in which production-when-required is probabilistic. This is a fact that does not accord well with the notion that the acquisition of grammar is a matter of the acquisition of rules, since the rules… either apply or do not apply. One would expect rule acquisition to be sudden. (Ref. [17], p. 257)

Is application of the regular past tense uniform?

Pinker stresses that symbolic rules do not vary in their applicability, but depend only on categorical conditions: the past tense applies to any verb stem. Does the evidence support the predicted uniformity? We consider four cases:

Uniformity with respect to phonology

Prasada and Pinker [19] tested judgments on and production of the past tense using nonce forms like *plip* or *ploamph*, manipulating phonological similarity to existing words. They concluded that there was an effect of similarity to known exceptions on novel irregular inflections, but no effect of similarity to known regulars for the regular inflection. However, there was an effect for regulars, which Prasada and Pinker attributed to a confound: their nonce stems, like *ploamph*, that were not similar to other regular items, were also phonologically strange. Even though subjects were asked to judge the inflection and not the stem, Prasada and Pinker claimed that the judgments were affected by the phonological properties of the stem, and 'corrected' for this by subtracting stem acceptability ratings. But this may be correcting away a real effect. A recent study by Albright and Hayes (unpublished manuscript) avoided the confound by using nonce stems of high phonological acceptability, and varied whether the item occurred in an 'island of reliability' for the regular or for an exceptional past tense. For example, their corpus contained over 300 verbs ending in an unvoiced fricative (e.g. *rush* or *laugh*); this is an island of reliability in that every such verb is regular. Both regular and irregular inflections received higher ratings if they came from reliable islands. The effect for regulars survived partialling out any competing influence favoring exceptions. Thus the regular past tense is sensitive to phonological attributes of the stem, violating the prediction of the symbolic rule account.

Uniformity with respect to semantics

A role for word meaning informing the regular past tense is vigorously rejected in Pinker's theory, because sensitivity to semantic similarity runs counter to the claimed encapsulation of the system that applies phonological transformations to word forms. Yet an influence of meaning in the selection of regular as well as irregular past-tense forms has often been argued [20–22]. In a recent study, Ramscar [22] placed nonce verbs like *frink* into semantic contexts that encouraged an interpretation resembling either *drink* or *blink*. The former typically elicited *frank* whereas the latter increased the likelihood of *frinked* (see Fig. 2). Contrary to Pinker's claims that denominal status blocks access to exceptions, a high level of *frank* responses occured even when subjects treated *frink* as denominal. Other experiments in Ramscar's study [22] demonstrated strong effects of contextually-specified meanings on inflection of *fly* as *flew* or *flied*, and again denominal status failed to block the choice of irregular *flew*. These findings clearly show that meaning can influence choice of the regular vs. irregular inflection, and fail to support the claim [5,23] that denominal status blocks access to lexically marked exceptions.

Semantic influences during acquisition

Shirai and Anderson [24] examined the use of the past tense as a function of semantic properties of the situation referred to in children's speech. When it first appears, the use of the past tense (including over-regularization) is largely restricted to descriptions of punctate events that have endpoints and produce results (such as '*I dropped it*'); it then gradually spreads to cases in which one of the typical properties (is punctate, has endpoint, produces results) is violated. The children's initial usage corresponds to the typical, but certainly not the only, cases that appear in their mother's speech, suggesting that initial use of the regular past grows from a semantic prototype.

The exception that proves the rule?

In English, the regular past is common, applying to 86% of the 1000 most common verbs [5]. Pinker [5,6] and Marcus *et al.* [25] have suggested, however, that occurrence in a high percentage of the verbs in a language is not necessary for the discovery of a regular pattern. Three cases have received the bulk of this discussion: (1) the regular German past participle +*t* [26]; (2) the Arabic broken plural [27]; and (3) the German +*s* plural [25]. Careful scrutiny of cases (1) and (2) [28,29] indicates that the forms in question may not be in the minority. So the case for 'the exception that proves the rule' [25] falls to the German +*s* plural. Marcus *et al.* claim that the +*s* plural, despite occurring in only a small fraction of German nouns, is the default used by German speakers whenever there is a 'failure of lexical memory'. They enumerate 21 separate contexts in which they suppose that lexical memory will fail, and argue that the +*s* plural should be used in all of these cases because it functions as a symbolic rule independent of the particular characteristics of the item to which it applies.

The +*s* plural certainly is in the minority in German. But does it apply uniformly as the symbolic rule account predicts? In fact, its usage is not uniform even in the Marcus *et al.* paper [25], which examined assignment of the +*s* plural to nonce forms treated as (a) unknown but real German words, (b) foreign words, or (c) proper names. For both (b) and (c) only the default rule should be available, and yet these two cases do not reveal the same pattern of extension of the +*s* plural. Hahn and Nakisa [30] (see Fig. 3) disconfirm the claim that +*s* acts uniformly across several of the contexts claimed by Marcus *et al.* The only case of high and nearly uniform use of +*s* occurs with surnames and does not extend fully even to first names: two members of the *Mann* family are called *Manns* but two girls named *Ulrike* can be two *Ulriken*. Bybee also notes relatively high probability for foreign borrowings ending in full vowels [26]. Surnamehood is an arbitrary property that must be associated with a specific use of an item in context, and assigning +*s* to foreign borrowings ending in full vowels requires

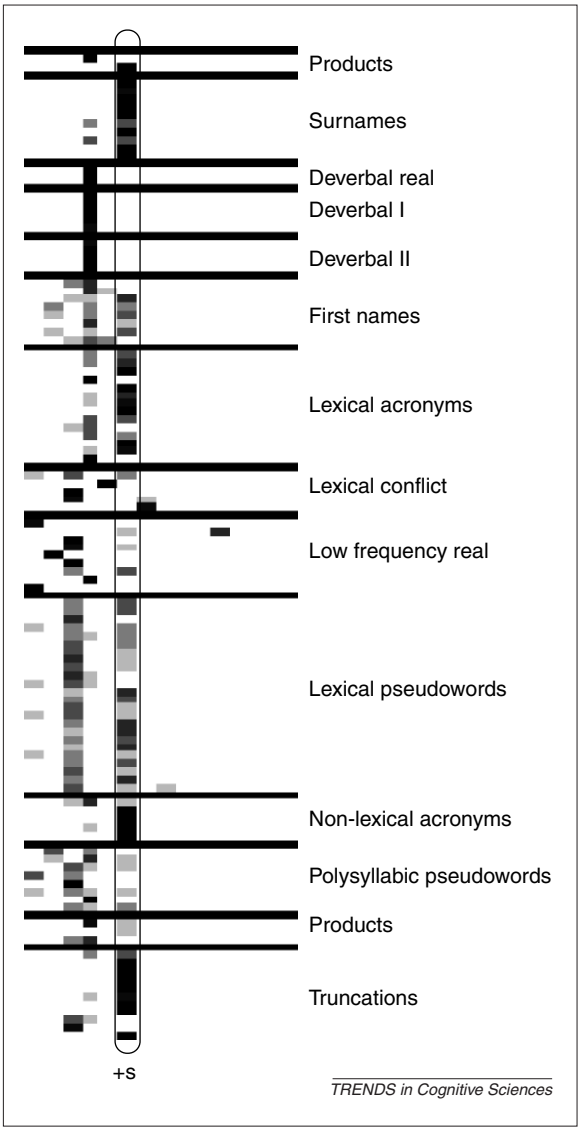

Fig. 3. Evidence that the German *+s* plural is not used uniformly across several situations supposedly calling for the use of a default as proposed by Marcus *et al.* [25]. Each row of the figure represents a different noun form, with the type of the form indicated; the horizontal bars separate the different types. Columns of the figure indicate alternative possible plural inflections, with the *+s* plural specifically highlighted. Grayscale darkness of the entry in each cell indicates the likelihood of using the particular plural for the given item, based on data from native German speaking adults. Reprinted with permission from Ref. [30].

sensitivity to phonology and etymology. Such specificity undercuts the notion that the German +*s* plural is in any sense a default. It is not the exception that proves the rule; instead it is another case with the graded, probabilistic and context-sensitive characteristics seen in connectionist networks.

Is regular inflection separable from inflection of exceptions?

Is there a separate mechanism for regular inflections? In contrast to the connectionist approach, the dual-mechanism theory argues that there is, and predicts the occurrence of selective deficits in producing and comprehending regular inflections. Pinker considered two putative examples [4]:

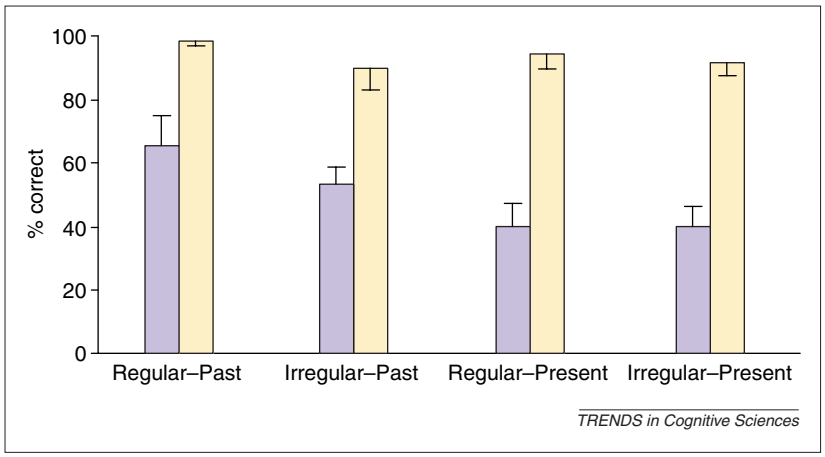

Fig. 4. Performance in an elicitation task requiring production of the past tense given the present (‘*Every day I wash my clothes; yesterday I ___ my clothes.*’) or of the present given the past (‘*Yesterday I washed my clothes; every day I ___ my clothes*’) for affected (mauve bars) and unaffected (yellow bars) members of the KE family. Results are based on matched sets of 10 regular and irregular verbs. Redrawn with permission from Ref. [35].

Genetic knockouts?

A large family (the KE family) consists of some normal individuals and some with an identified single-gene defect [31,32]. Reports based on testing with a small number of stimuli [33,34] suggested that affected individuals had special difficulty with regular compared with irregular inflections. Subsequent investigation by Vargha-Khadem *et al.* [35], however, painted a different picture. Affected family members were found to have a wide range of deficits in linguistic and non-linguistic tasks, and they demonstrated substantial and equal difficulty with regular and irregular forms (Fig. 4) when tested with a longer and better-controlled list. There was no sign of selective vulnerability of the regular inflection. We do not rule out the possibility that a developmental phonological deficit could result in difficulty acquiring regular forms [36]. Indeed, if regular inflections are phonetically weak in the input to a network, an impairment in phonological representation can result in a failure to learn the regular past tense [37]. This provides one way of understanding why some children diagnosed with specific language impairment present with an apparent selective deficit in inflectional morphology and other aspects of grammar [38], as many aspects of grammar are signalled by phonetically weak material [39].

Effects of brain damage?

Anterior lesions in the left hemisphere often result in dysfluent speech containing few grammatical morphemes or inflections [40]. Ullman *et al.* [41,42] have reported a patient of this type who produced the correct past tense for 69% of exceptions but only 20% of regulars and 5% of nonce forms in a past-tense elicitation task. In collaboration with several others [43] we have considered the possibility that an uncontrolled difference between the regular and exception items in Ullman’s study could have influenced the results: the word-final consonant clusters were longer, on average, in the regular past tenses (2.0 consonants) than in the exceptions (1.2 consonants). This is natural, because regular inflection involves the addition of phonological material to the verb stem, thereby increasing its complexity [44]. By contrast, the formation of exceptions generally involves a vowel and/or consonant change (*eat–ate*, *think–thought*) that tends to conserve complexity. Where something is added, there is typically a compensatory reduction in vowel length (*keep–kept*), so that exceptional past tenses fall within acceptable phonological bounds.

Bird *et al.* [43] identified 10 non-fluent aphasic patients who were all significantly better with irregular verbs on a screening list unmatched for phonological factors. The advantage occurred in the elicitation task (37% vs. 20% correct), and also in single-word repetition (68% vs. 47%) and single-word reading (44% vs. 24%). When tested with regular and exception past tenses matched for phonological complexity, the patients no longer showed an advantage for irregulars in the elicitation task (means of 26% irregular, 29% regular) or in repetition (65% irregular vs. 64% regular), supporting the view that the initial difference was phonological rather than morphological in origin. A remaining irregular advantage in reading (41% vs. 27%) was interpreted as a concreteness effect: past-tense verbs like *ground* and *rose* are also concrete nouns.

Ullman *et al.* [41] also reported a disadvantage in the elicitation task for regular verbs in patients with Parkinson’s Disease (PD). Again, however, the effect can be interpreted in terms of phonological complexity because, in the specially designed ‘PD retest’ list, onset consonant clusters were longer in the regular than the irregular verbs. Furthermore, the disadvantage reported for non-words relative to exceptions cannot be attributed to inflectional processes: the PD patients’ responses to non-words, although often characterized by stem distortions (*pragged* or *planned* instead of *plagged*), were correctly inflected 91% of the time (vs. 88% for the exceptions).

Summary of the state of the evidence

In Table 1 we listed contrasting predictions of the dual-mechanism and PDP theories. Our review of the evidence suggests that the onset of the regular past (and all other inflections) is gradual rather than sudden; that both the English regular past tense and the German +*s* plural are subject to phonological, semantic and other influences rather than being uniform in their application; and that there is no convincing evidence that the inflection of regular verbs can be selectively impaired, except insofar as such impairment is a direct or indirect consequence of a phonological impairment. The evidence seems therefore to be fully compatible with the idea that

inflectional processes arise in a single integrated system, in which graded and context-sensitive influences of many different types jointly determine whether a regular or an exceptional past tense (or other inflection) will apply. This single system has all of the characteristics of the connectionist framework for inflectional processing.

We do not claim that it would be impossible to construct a rule-based model of inflection formation that has all of the properties supported by the evidence. However, such an account would not be an instantiation of Pinker's symbolic rule account. In fact, rule-based models with some of the right characteristics are currently being pursued ([45]; Albright and Hayes, unpublished). If such models use graded rule activations and probabilistic outcomes, allow rules to strengthen gradually with experience, incorporate semantic and phonological constraints, and use rules within a mechanism that also incorporates word-specific information, they could become empirically indistinguishable from a connectionist account. Such models might be viewed as characterizing an underlyingly connectionist processing system at a higher level of analysis, with rules providing descriptive summaries of the regularities captured in the network's connections.

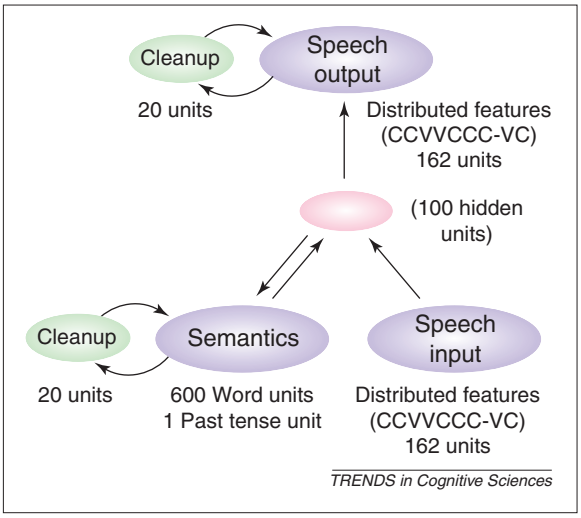

Fig. 5. The connectionist model of Joanisse and Seidenberg [46], in which regular and inrregular forms are generated by a single system, using phonological input and output representations and a semantic internal representation. When a verb is presented on the input, the network is trained to generate an appropriate semantic representation (activating the correct word unit and the past tense unit if appropriate) and also to generate the corresponding output representation. The network is also trained to produce the corresponding phonological output when given an input activating anindividual semantic unit corresponding to each taught word, and to generate past tenses when the past tense unit is activated and either a verb stem is presented to the phonological input or a word unit is activated in semantics. Redrawn with permission from Ref. [46].

Towards an adequate connectionist account

Existing connectionist models still have limitations. Given the extent of empirical support for the predictions arising from the connectionist approach, however, we remain convinced of the fruitfulness of pursuing the approach. Our current efforts build on a model by Joanisse and Seidenberg [46] (Fig. 5), which incorporates a role for semantic representations (see also Refs [13,14]), something left out of Rumelhart and McClelland's original formulation [8] as a simplification. This model can explain why a semantic deficit disproportionately disrupts production of exceptional past tenses, as demonstrated by Ullman *et al.* [41,42] and Patterson *et al.* [47]: word meaning provides information that helps the network to treat a particular item distinctively, counteracting the network's tendency to apply the regular inflection. Some limitations remain, however. Our extensions will use distributed semantic representations that capture similarity in meaning, as well as refinements to phonological processes to address phonological complexity and perceptibility effects. The fact that such a complete model is not yet implemented is scarcely surprising or unique. Encompassing the whole problem is a real challenge for any model, and current rule-based proposals are at best only partially implemented.

In pointing towards a future connectionist account, we note one significant aspect that might be under-appreciated. Contrary to some statements (e.g. Ref. [4]), connectionist networks are not simply analogy mechanisms that base their tendency to generalize on raw item-to-item similarity [48]. Instead, they are sensitive to regularities, so that if an input–output relationship is fully regular, the network can closely approximate a categorical, symbolic rule. Such a property is necessary if these models are to capture the full range of inflectional systems, because there are cases throughout the world's languages (including the English progressive, *-ing*, form) that are completely regular [49]. These occur among many other cases with varying degrees of regularity, and networks of the right sort should be able to capture the whole spectrum. This makes the connectionist network fundamentally different from either the symbolic rule or the lexical mechanism considered in the dual-mechanism account.

Acknowledgements
Preparation of this article was supported by MH 47566, MH 64445 and by the Medical Research Council. We thank David Plaut, Michael Ramscar and Mark Seidenberg for discussion and comments. We acknowledge Pinker and Ullman's generous comments about the role of the Rumelhart–McClelland model in stimulating psycholinguistic research. Clearly, though, just as much of the credit goes to Stephen Pinker and his collaborators, including Michael Ullman, for taking up the challenge posed by the RM model in this vigorous ongoing debate.

References

1 Chomsky, N. (1957). *Syntactic Structures*, Mouton
2 Chomsky, N. (1959) Review of Skinner (1957). *Language* 35, 26–58
3 Fodor, J. and Pylyshyn, Z. (1988) Connectionism and cognitive architecture: a critical analysis. *Cognition* 28, 3–71
4 Pinker, S. (1991) Rules of language. *Science* 253, 530–535
5 Pinker, S. (1999) *Words and Rules: The Ingredients of Language*, HarperCollins
6 Pinker, S. (2001) Four decades of rules and associations, or whatever happened to the past tense debate? In *Language, The Brain, and Cognitive Development: Papers in Honor of Jacques Mehler*. (Dupoux, E. ed.), MIT Press
7 Pinker, S. and Prince, A. (1988) On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193
8 Rumelhart, D.E. and McClelland, J.L. (1986) On learning the past tenses of English verbs. In *Parallel Distributed Processing* (*Vol 2*): *Psychological and Biological Models* (McClelland J.L. and Rumelhart D.E., eds), pp. 216–271, MIT Press
9 McClelland, J.L. *et al.* (1986) The appeal of parallel distributed processing. In *Parallel*

Distributed Processing (*Vol.1*): *Explorations in the Microstructure of Cognition* (Rumelhart, D.E. and McClelland, J.L., eds), pp.3–44, MIT Press
10 Pinker, S. and Ullman, M. (2002) The past and future of the past tense. *Trends Cogn. Sci.* 6, 456–463
11 Ervin, S.M. (1964) Imitation and structural change in children's language. In *New Directions in the Study of Language* (Lenneberg, E. ed.), MIT Press
12 Berko, J. (1958) The child's learning of English morphology. *Word* 14, 150–177
13 Seidenberg, M.S. and McClelland, J.L. (1989) A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 96, 523–568
14 Plaut, D.C. *et al.* (1996) Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol. Rev.* 103, 56–115
15 Marcus, G.F. *et al.* (1992) Overregularization in language acquisition. *Monogr. Soc. Res. Child Dev.*, Serial No. 228, Vol 57, No. 4
16 Cazden, C.B. (1968) The acquisition of noun and verb inflections. *Child Dev.* 39, 433–448
17 Brown, R. (1973) *A First Language*, Harvard University Press
18 MacWhinney, B. and Snow, C. (1990) The child language data exchange system: an update. *J. Child Lang.* 17, 457–472
19 Prasada, S. and Pinker, S. (1993) Generalization of regular and irregular morphological patterns. *Lang. Cogn. Process.* 8, 1–56
20 Harris, C.L. (1992) Understanding the English past-tense formation: the shared meaning hypothesis. In *Proc. 14th Annu. Conf. Cogni. Sci. Soc.*, Erlbaum
21 Daugherty, K. *et al.* (1993) Why no mere mortal has ever flown out to center field but people often say they do. In *Proc. 15th Annu. Conf. Cogn. Sci. Soc.*, pp. 383–388, Erlbaum
22 Ramscar, M. (2002) The role of meaning in inflection: why the past tense doesn't require a rule. *Cogn. Psychol.* 45, 45–94
23 Kim, J.J. *et al.* (1991) Why no mere mortal has ever flown out to center field. *Cogn. Sci.* 15, 173–218
24 Shirai, Y. and Anderson, R.W. (1995) The acquisition of tense-aspect morphology: a prototype account. *Language* 71, 743–762
25 Marcus, G. *et al.* (1995) German inflection: The exception that proves the rule. *Cogn. Psychol.* 29, 189–256
26 Clahsen, H. and Rothweiler, M. (1992) Inflectional rules in children's grammars: evidence from the development of participles in German. *Yearbook Morphol.* 1, 1–34
27 McCarthy, J. and Prince, A. (1990) Foot and word in prosodic morphology: the Arabic broken plural. *Nat. Lang. Linguist. Theory* 8, 209–283
28 Bybee, J.L. (1995) Regular morphology and the lexicon. *Lang. Cogn. Process.* 10, 425–455
29 Boudelaa, S. and Gaskell, M.G. (2002) A re-examination of the default system for Arabic plurals. *Lang. Cogn. Process.* 17, 321–343
30 Hahn, U. and Nakisa, R.C. (2000) German inflection: single-route or dual-route? *Cogn. Psychol.* 41, 313–360
31 Vargha-Khadem, F. *et al.* (1998) Neural basis of an inherited speech and language disorder. *Proc. Natl. Acad. Sci. U. S. A.* 95, 12695–12700
32 Lai, C.S.L. *et al.* (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413, 519–523
33 Gopnik, M. (1990) Dysphasia in an extended family. *Nature* 344, 715
34 Gopnik, M. and Crago, M. (1991) Familial aggregation of a developmental language disorder. *Cognition* 39, 1–50
35 Vargha-Khadem, F. *et al.* (1995) Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc. Natl. Acad. Sci. U. S. A.* 92, 930–933
36 Leonard, L. (1998) *Children with Specific Language Impairment*, MIT Press
37 Hoeffner, J. and McClelland, J.L. (1993) Can a perceptual processing deficit explain the impairment of inflectional morphology in developmental dysphasia? A computational investigation. In *Proc. 25th Annu. Child Lang. Res. Forum* (Clark, E., ed.), pp. 38–49, Center for the Study of Language and Information, Stanford University
38 van der Lely, H.K. *et al.* (1998) Evidence for a grammar-specific deficit in children. *Curr. Biol.* 8, 1253–1258
39 Leonard, L.B. *et al.* (1992) Grammatical morphology and speech perception in children with specific language impairment. *J. Speech Hear. Res.* 35, 1076–1085
40 Marin, O. *et al.* (1976) Dissociations of language in aphasia: implications for normal function. *Ann. New York Acad. Sci.* 280, 864–884
41 Ullman, M.T. *et al.* (1997) A neural dissociation within language: evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *J. Cogn. Neurosci.* 9, 266–276
42 Ullman, M.T. *et al.* Neural correlates of lexicon and grammar: evidence from the production, reading, and judgment of inflection in aphasia. *Brain Lang.* (in press)
43 Bird, H. *et al.* Deficits in phonology and past tense morphology: what is the connection? *J. Mem. Lang.* (in press)
44 Burzio, L. (2002) Missing players: phonology and the past tense debate. *Lingua* 112, 157–199
45 Taatgen, N. and Anderson, J. Why do children learn to say 'broke'? A model of learning the past tense without feedback. *Cognition* (in press)
46 Joanisse, M.F. and Seidenberg, M.S. (1999) Impairments in verb morphology following brain injury: a connectionist model. *Proc. Natl. Acad. Sci. U. S. A.* 96, 7592–7597
47 Patterson, K. *et al.* (2001) Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia* 39, 709–724
48 Plaut, D.C. and Gonnerman, L.M. (2000) Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Lang. Cogn. Process.* 15, 445–485
49 Dabrowska, E. (2001) Learning a morphological system without a default: the Polish genitive. *J. Child Lang.* 28, 545–574

Combination and structure, not gradedness, is the issue

Reply to McClelland and Patterson

McClelland and Patterson's Opinion article [1] largely hinges on whether the regular past tense is acquired instantaneously and applied perfectly, which they consider to be hallmarks of symbolic models. McClelland and Patterson take gradedness in behavioral data as evidence for the connectionist approach. We believe this framing sidesteps the key issue in the past-tense debate: whether human language uses mechanisms that are combinatorial and sensitive to grammatical structure and categories.

Symbolic models of cognition [2] and our approach to language in particular (see [3] Chap. 5; and [4] pp. 130–136) have always invoked combinatorial operations ('rules') that are acquired gradually and can be applied probabilistically. Less-than-100% application of a regular inflection can occur for many reasons: intermediate stages in acquisition, partial blocking by weak irregulars, phonotactic naturalness, depth of processing of the grammatical structure, uncertainty as to whether a rule's conditions have been met, and the noisiness of neural computation. An absence of step-functions or all-or-none data is thus questionable evidence for connectionism.

 1364-6613/02/$ – see front matter PII: S1364-6613(02)02013-2

More germane is whether regular inflection is always available to generate an acceptable form when memory fails, whether it applies in heterogeneous circumstances whose only common denominator is the word's grammatical category, and whether it neuropsychologically dissociates from memory lookup and associates with combinatorial processing.

Acquisition

McClelland and Patterson argue that acquisition of regular tense-marking is not a step-function, but we never claimed it was[a]. The analysis they dispute only supported the uncontroversial idea that the English past-tense is not innate and that application of the suffix to regular and (sometimes) irregular verbs should develop in tandem [5]. This idea, together with the possibility that children can store unanalyzed words, is sufficient to explain 'U-shaped' development of irregulars; the connectionist prediction that over-regularization is triggered by a sudden increase in regular forms in the input is both empirically incorrect and theoretically unnecessary [5–7].

Systematic regularization

Ramscar's claim that this phenomenon (*rang the bell/ringed the city*) can be reduced to semantic dissimilarity is incompatible with the distribution of regular/irregular homophones in English: virtually no polysemous irregular roots tie regular forms to specific meanings (**throwed up*) unless they are exocentric, and virtually all exocentric irregular-sounding forms are regularized [8–10]. (Thus even Joanisse and Seidenberg conceded that semantic similarity is 'not important for the past tense.') Ramscar's experiment used a single, unrepresentative item, confounded lexical with semantic differences, and was tainted by demand characteristics: people were in effect given the question 'Does the experimenter want me to treat *frink* as a distorted version of *drink*, or of *blink*?' Ramscar's intended manipulation of exocentric structure was ineffective because it used odd semantic relationships found in no English verb, and the cursory presentation gave participants no inducement to take it seriously.

German inflection

We never conceded that German *-t* participles are irrelevant to the connectionist hypothesis about the hallmarks of regularity, namely that they are an epiphenomenon of regular forms constituting the 'overwhelming majority' of the child's input [11,12]. Our claim was that even if one bent over backwards and recounted words using criteria maximally unfavorable to our position, the German *-s* plural would disprove the hypothesis. But we don't accept the criteria. Counts that put *-t* in the majority require collapsing morphologically related non-compositional words (although connectionism eschews morphological structure), counting types (although connectionist models are driven by tokens, for which regulars are not in the majority by any criteria, even in English), and using huge corpora containing many obscure words.

We agree that the uneven applicability of *-s* to the different default circumstances in German requires additional explanation (see [12]). But the data are more poorly explained by McClelland and Patterson's alternative that German speakers learn to connect *-s* with each 'arbitrary property that must be associated with a specific use of an item in context', such as surnamehood[b]. This leaves it a coincidence that the circumstances eliciting *-s* (names, unassimilated borrowings, unusual-sounding words, acronyms, truncations, quotations, onomatopoeia, nominalized phrases and conjunctions) all involve failure to access an irregular root but have nothing in common semantically or phonologically [11–13]. It also does not explain why speakers use *-s* in circumstances too rare for them to have been trained on beforehand (e.g. quotations, as in the German equivalent of 'I found three *man*'s on page 1').

Genetic impairments

Although we once cited a preliminary finding that in Specific Language Impairment (SLI), regulars are harder than irregulars (calling for the same explanation as for agrammatism) [14], our own and other subsequent analyses show no difference [15–20]. The best explanation is that language-impaired people are indeed impaired with rules (as seen in their poor performance when inflecting nonsense words) but can memorize common regular forms (hence the lack of a deficit compared with irregulars)[15–17]. Supporting evidence is that regulars show consistent frequency effects in SLI but not in control subjects [15–19]. This suggests that children growing up with a grammatical deficit are better at compensating for it via memorization than are adults who acquired their deficit later in life.

McClelland and Patterson claim that pattern associators can explain a regular–irregular difference as a by-product of a deficit in processing unstressed material. However, such a difference does not exist, and the hypothesis that SLI is caused by a perceptual deficit is no longer tenable. Children can have SLI without auditory processing deficits and vice-versa, and people with SLI have trouble on grammatical tasks but not on phonologically matched control tasks [21–23].

Aphasia

Bird *et al.* [24] replicate eight earlier studies showing that non-fluent aphasics have more trouble with

Steven Pinker
Dept of Brain and Cognitive Sciences, NE20-413, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
e-mail: steve@psyche.mit.edu

Michael Ullman
Dept of Neuroscience, Research Building EP-04, Georgetown University, 3900 Reservoir Rd, NW, Washington DC 20007, USA.
e-mail: michael@georgetown.edu

[a]Terms like 'epiphany' and 'deduces' were used informally in Ref. [5] as a shorthand for the process by which children acquire the past tense. The context (pp. 202–203) explicitly discusses the gradual development and probabilistic application of the rule.

[b]The pluralized name *Ulrike/Ulriken* is not a counterexample, both because the *-en* plural strikes many speakers as archaic or jocular, and because the feminine suffix *-e* itself selects *-en* (see Ref. [11], Note 18; [12,35]).

regular than irregular forms in generation, reading, and repetition [25–33]. Most took measures to equate phonological complexity. Bird *et al*. implemented additional controls involving subsets of items or multiple regressions, and obtained mixed results. The regular–irregular difference disappeared in the new analyses of the generation task, survived in the reanalyses for the reading task, and disappeared in one analysis of a repetition task but survived in another. Further complicating this mixed picture is that Bird *et al*.'s irregular items had a greater complexity of stem-to-past mappings than in earlier studies, and their regular list included items that rhymed with irregulars (which are likely to be memorized [25,34], leaving them less vulnerable to the effects of agrammatism).

Bird *et al*.'s study comparing discrimination of regular stems and pasts (*press/pressed*) to discrimination of phonologically matched words (*chess/chest*) is also equivocal. Most patients were either at chance or ceiling at both tasks, and the others showed greater difficulty with the past-tense discrimination, which is consistent with other studies. While we applaud the extensive testing and careful design of the Bird *et al*. study, we believe they have not demonstrated that the regular–irregular difference in aphasia is an epiphenomenona of phonological complexity.

Connectionist models

We agree that connectionist networks are not always analogy mechanisms. Our point (based on explications by McClelland and other connectionists) is that pattern associators (the most common connectionist model of the past tense) tend towards analogy when learning competing patterns under standard training regimes. This is what gives such models their predictive power with irregular forms. The claim that some connectionist model can, given a specific architecture, training schedule and input features, approximate any linguistic phenomenon might be true, but it is in danger of reducing connectionism to a universal statistical approximation technique rather than a source of empirical predictions. Language cannot be treated as just a collection of 'regularities in the input' that can be approximated by some mechanism; those regularities are themselves the products of human minds and need to be explained.

References

1 McClelland, J.L. and Patterson, K. (2002) Rules or connections in past-tense inflections: what does the evidence rule out? *Trends Cogn. Sci.* 6, 465–472
2 Anderson, J.R. (1993) *Rules of the Mind*, Erlbaum
3 Pinker, S. (1984) *Language Learnability and Language Development*, Harvard University Press
4 Pinker, S. and Prince, A. (1988) On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193
5 Marcus, G.F. *et al*. (1992) Overregularization in language acquisition. *Monogr. Soc. Res. Child Dev.* 57 (Serial No. 228), 1–165
6 Marcus, G.F. (1995) Children's overregularization of English plurals: a quantitative analysis. *J. Child Lang.* 22, 447–459
7 Marcus, G.F. (1995) The acquisition of the English past tense in children and multilayered connectionist networks. *Cognition* 56, 271–279
8 Pinker, S. (1999) *Words and Rules: The Ingredients of Language*, Basic Books
9 Kim, J.J. *et al*. (1991) Why no mere mortal has ever flown out to center field. *Cogn. Sci.* 15, 173–218
10 Kim, J.J. *et al*. (1994) Sensitivity of children's inflection to morphological structure. *J. Child Lang.* 21, 173–209
11 Marcus, G.F. *et al*. (1995) German inflection: the exception that proves the rule. *Cogn. Psychol.* 29, 189–256
12 Clahsen, H. (1999) Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behav. Brain Sci.* 22, 991–1060
13 Clahsen, H. (1997) The representation of participles in the German mental lexicon: evidence for the dual-mechanism model. In *Yearbook of Morphology* (Booij, G. and van Marle, J., eds), pp. 73–95, Kluwer Academic Publishers
14 Pinker, S. (1991) Rules of language. *Science* 253, 530–535
15 Ullman, M.T. and Gopnik, M. (1994) The production of inflectional morphology in hereditary specific language impairment. In *The McGill Working Papers in Linguistics: Linguistic Aspects of Familial Language Impairment* (Vol. 10) (Matthews, J., ed.), pp. 81–118, McGill University
16 Ullman, M.T. and Gopnik, M. (1999) Inflectional morphology in a family with inherited specific language impairment. *Appl. Psycholinguist.* 20, 51–117
17 van der Lely, H.K.J. and Ullman, M.T. (2001) Past tense morphology in specifically language impaired and normally developing children. *Lang. Cogn. Processes* 16, 177–217
18 Oetting, J.B. and Rice, M. (1993) Plural acquisition in children with specific language impairment. *J. Speech Hear. Res.* 36, 1236–1248
19 Oetting, J.B. and Horohov, J.E. (1997) Past tense marking by children with and without specific language impairment. *J. Speech Hear. Res.* 40, 62–74
20 Vargha-Khadem, F. *et al*. (1995) Praxic and nonverbal cognitive deficits in a large family with genetically transmitted speech and language disorder. *Proc. Natl. Acad. Sci. U. S. A.* 92, 930–933
21 van der Lely, H.K.J. *et al*. (1998) Evidence for a grammar-specific deficit in children. *Curr. Biol.* 8, 1253–1258
22 Briscoe, J. *et al*. (2001) Phonological processing, language and literacy: a comparison of children with mild to moderate sensorineural hearing loss and those with specific language impairment. *J. Child Psychol. Psychiatry* 42, 329–340
23 Bishop, D.V.M. *et al*. (1999) Auditory temporal processing impairment: neither necessary nor sufficient for causing language impairment in children. *J. Speech Lang. Hear. Res.* 42, 1295–1310
24 Bird, H. *et al*. Deficits in phonology and past tense morphology: what is the connection? (in press)
25 Ullman, M.T. *et al*. (in press) Neural correlates of lexicon and grammar: evidence from the production, reading, and judgment of inflection in aphasia. *Brain Lang.*
26 Ullman, M. *et al*. (1997) A neural dissociation within language: evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *J. Cogn. Neurosci.* 9, 289–299
27 Coslett, H.B. (1986) Dissociation between reading of derivational and inflectional suffixes in two phonological dyslexics. Presented at *Academy of Aphasia*, Nashville, Tennessee (October 1986)
28 Coslett, H.B. (1988) A selective morphologic impairment in writing: evidence from a phonological dysgraphic. Presented at *Academy of Aphasia*, Montreal, Canada (October 1988)
29 Marin, O. *et al*. (1976) Dissociations of language in aphasia: Implications for normal function. In *Origin and Evolution of Language and Speech* (Vol. 280) (Harnad, S.R. *et al*., eds), New York Academy of Sciences
30 Badecker, W. *et al*. (1991) Varieties of sentence comprehension deficits: a case study. *Cortex* 27, 311–321
31 Badecker, W. and Caramazza, A. (1987) The analysis and morphological errors in a case of acquired dyslexia. *Brain Lang.* 32, 278–305
32 Penke, M. and Krause, M. (1999) Broca's aphasia and German plural formulation. *Brain Lang.* 69, 305–318
33 Hagiwara, H. *et al*. (1999) Neurolinguistic evidence for rule-based nominal suffixation. *Language* 75, 739–763
34 Ullman, M.T. (2001) The declarative/procedural model of lexicon and grammar. *J. Psycholinguist. Res.* 30, 37–69
35 Wiese, R. (1986) Shwa and the structure of words in German? *Linguistics* 24, 695–724