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* ___',

http://tics.trends.com 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].

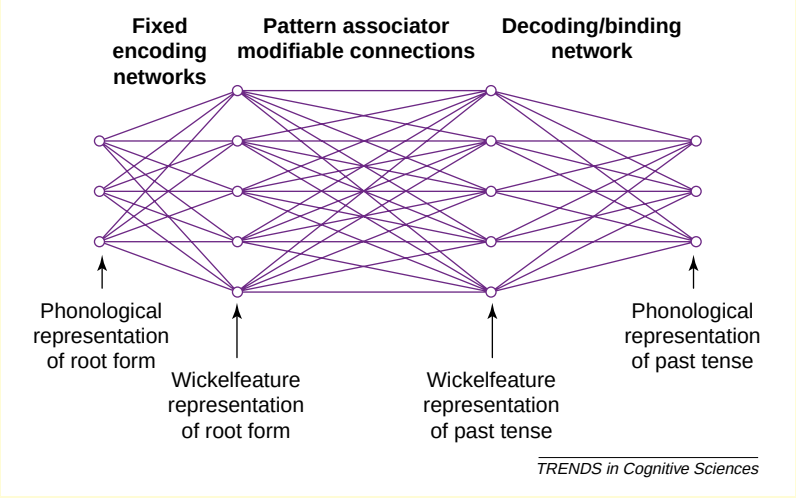

Fig. I. The Rumelhart–McClelland model of past-tense inflection (see text for discussion). Reprinted with permission from Ref. [a].

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.

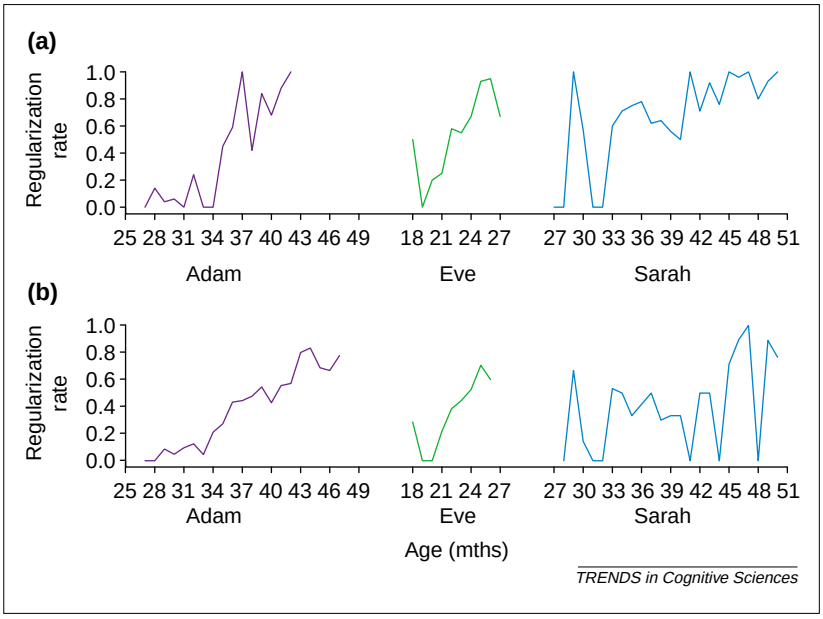

Fig. 1. Acquisition of the regular past tense by three children, Adam, Eve, and Sarah, as presented in Marcus *et al.* [15] and in Hoeffner (PhD thesis, Carnegie Mellon University, 1996). Percent usage of the regular past tense in obligatory contexts is plotted as a function of the child's age in months. (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

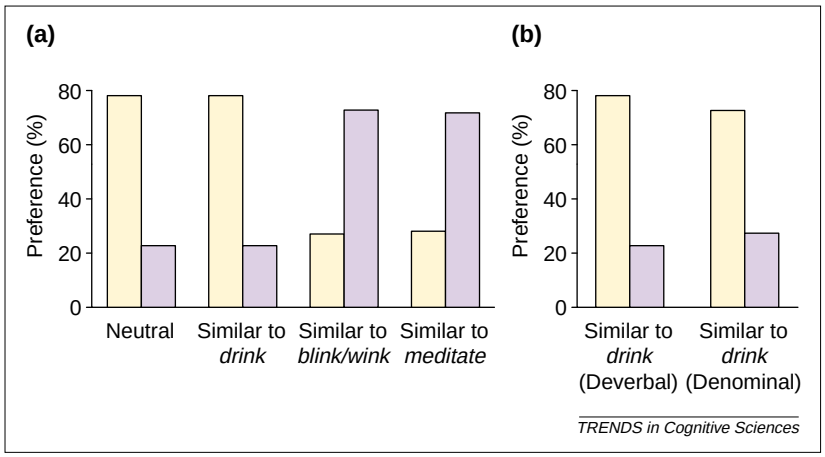

Fig. 2. Summary of effects of semantics and grammar on inflections of the nonce verbs *frink* and *sprink* from Ramscar [22]. (a) Use of irregular (*frank* or *sprank*, yellow bars) or regular (*frinked* or *sprinked*, mauve bars) in four different conditions. Note that in a neutral condition, with no 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

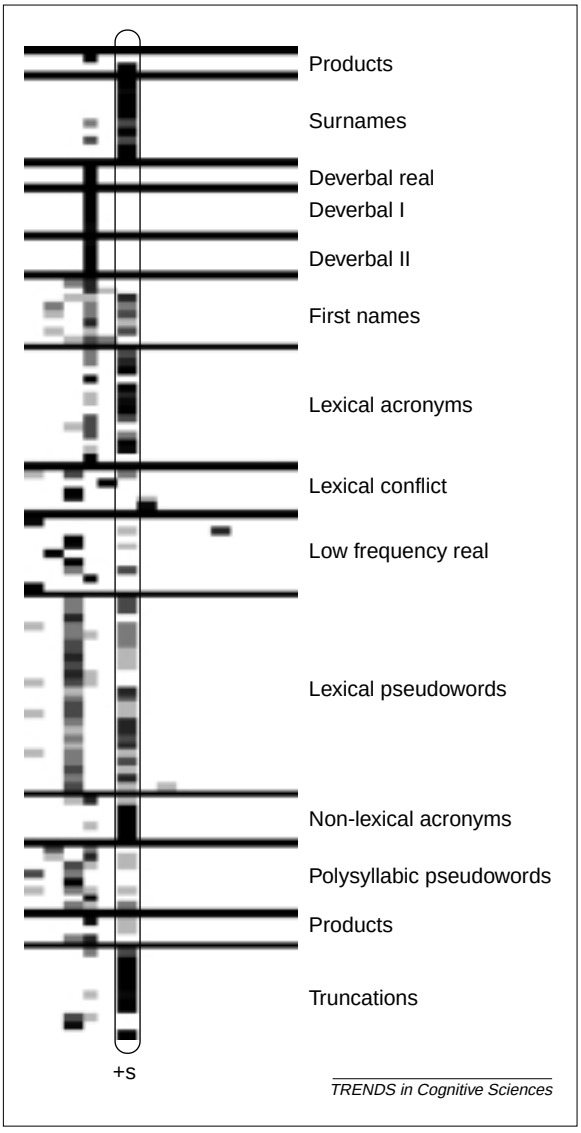

Fig. 3. Evidence that the German *+s* plural is not used uniformly across several situations supposedly calling for the use of a default as proposed by Marcus *et al.* [25]. Each row of the figure represents a different noun form, with the type of the form indicated; the horizontal bars separate the 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]:

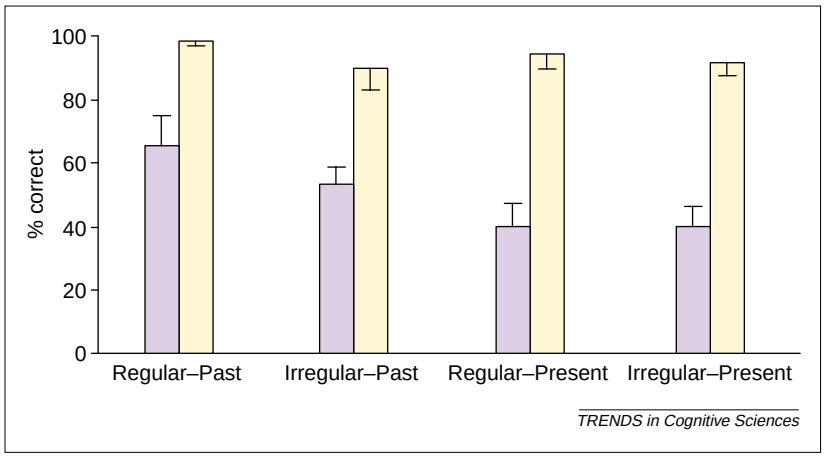

Fig. 4. Performance in an elicitation task requiring production of the past tense given the present ('*Every day I wash my clothes; yesterday I ___ my clothes.*') or of the present given the past ('*Yesterday I washed my clothes; every day I ___ my clothes*') for affected (mauve bars) and unaffected (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.

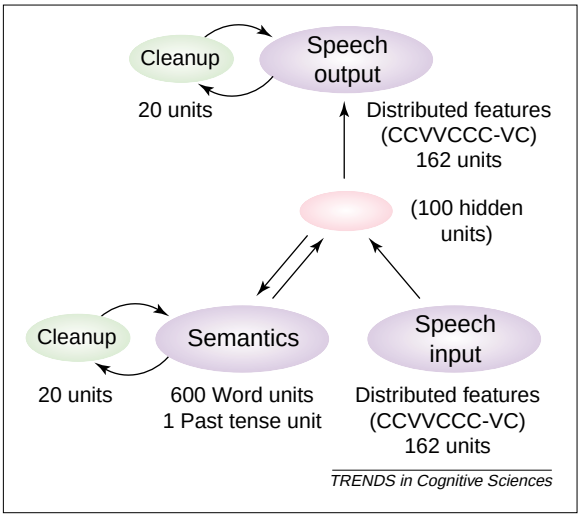

Fig. 5. The connectionist model of Joanisse and Seidenberg [46], in which regular and inrregular forms are generated by a single system, using phonological input and output representations and a semantic internal representation. When a verb is presented on the input, the network is trained to generate 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