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
10 Harris, C.L. (1994) Coarse coding and the lexicon. In Continuity in Linguistic Semantics (Fuchs, C. and Victorri, B., eds), J. John Benjamins

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 fell [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 ____’,
the response is usually ricked. As the child would never have heard go 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

**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. 1). 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].

**References**

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
opinion

Redrawn with permission from Ref. [22].

ratings were not affected by their judgment of whether the nonce verb seemed to be denominal.

tense, suggesting that use of the regular past tense can be influenced by semantics. (b) Subjects’ context, participants preferred irregular past tenses, and this trend persisted when context provided a sprink from Ramscar [22]. (a) Use of irregular (frank or sprank, yellow bars) or regular (frinked or

Fig. 2.

Preference (%) Preference (%)
0 0
20 20
40 40
60 60
80 80
Neutral Similar to drink Similar to blink/wink Similar to meditate
Similar to drink (Deverbal) Similar to drink (Denominal)

TRENDS in Cognitive Sciences

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 occurred 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.

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:

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 the 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)

http://tics.trends.com
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] fails 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 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. 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].
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.

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 underappreciated. 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.

References
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.

---


---