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 unnaturalness, 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.
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. 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/ ranged the tidy) 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 (*threw up) unless they are exocentric, and virtually all exocentric irregular-sounding forms are regularized [8–10]. (Thus even J. vanisse 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. 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), irregulars are harder than irregulars (calling for the same explanation as forgrammatism) [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
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 ofagrammatism).

Bird et al.'s study comparing discrimination of regular stems and pasts (press/presed) 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 epiphenomenon of phonological complexity.

**Connectionist models**

We agree that connectionist networks are not always analogy mechanisms. Our point (based on explanations 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**