COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 38 (2014) 1190–1228

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12147

Quasiregularity and Its Discontents: The Legacy of the Past Tense Debate

Mark S. Seidenberg,[a] David C. Plaut[b]

[a]*Department of Psychology, University of Wisconsin-Madison*
[b]*Department of Psychology, Carnegie Mellon University*

Received 20 June 2012; received in revised form 9 September 2013; accepted 4 November 2013

Abstract

Rumelhart and McClelland's chapter about learning the past tense created a degree of controversy extraordinary even in the adversarial culture of modern science. It also stimulated a vast amount of research that advanced the understanding of the past tense, inflectional morphology in English and other languages, the nature of linguistic representations, relations between language and other phenomena such as reading and object recognition, the properties of artificial neural networks, and other topics. We examine the impact of the Rumelhart and McClelland model with the benefit of 25 years of hindsight. It is not clear who "won" the debate. It is clear, however, that the core ideas that the model instantiated have been assimilated into many areas in the study of language, changing the focus of research from abstract characterizations of linguistic competence to an emphasis on the role of the statistical structure of language in acquisition and processing.

Keywords: Past tense debate; PDP models; Quasiregularity; Words and rules

1. Introduction

Chapter 18 in the second volume of the Parallel Distributed Processing (PDP) books created a fuss. The title, "On learning the past tenses of verbs in English," was one that, like "The information available in brief visual presentations" (Sperling, 1960) or "Studies of interference in serial verbal reactions" (Stroop, 1935) gave little hint as to its significance. The chapter described the application of some of the tenets of the PDP approach to a sliver of English grammar: the inflection of verbs for the past tense. Why the past tense? Because it exhibits three interesting characteristics. First, it is systematic: Most past tenses are formed by adding the morpheme that is spelled *-ed* and pronounced as in the examples *baked*, *baited*, and *bared*. Second, it is productive: People can readily

Correspondence should be sent to Mark Seidenberg, Department of Psychology, 1202 West Johnson Street, University of Wisconsin-Madison, Madison, WI 53706. E-mail: seidenberg@wisc.edu

generate past tenses for novel forms such as *nust-nusted* or *wug-wugged*. Third, it is quasiregular (Seidenberg & McClelland, 1989): There is a main pattern but also irregular forms that deviate from it in differing degrees (e.g., *keep-kept*, *run-ran*, *go-went*). Phenomena such as tense on verbs and number on nouns have been taken as simple, decisive demonstrations that grammatical rules are an essential component of linguistic knowledge (Pinker, 1999). Irregular forms exist outside this system of core linguistic knowledge and are learned and generated by other mechanisms such as memorization and association.

Rumelhart and McClelland (1986; hereafter RM) described a neural network model that presented an alternative to this linguistic orthodoxy. The model performed a version of the WUG task (Berko, 1958): It took the phonological form of a verb's present tense as input and generated its phonologically specified past tense as output. The "uniform procedure" by which the model generated past tenses obviated the distinction between rule-governed forms and exceptions. Moreover, the model did not simply memorize the patterns on which it had been trained; knowledge was represented in a form that supported the generation of past tense forms of novel verbs.

The chapter generated an enormous amount of controversy, acting as ground zero for critiques of the PDP approach. It is surely one of the most closely analyzed and criticized documents in cognitive science, starting with three long critiques in an issue of the journal *Cognition*, the most influential being Pinker and Prince's (1988). The chapter also launched a prodigious amount of research, on the past tense, on related phenomena in English and other languages, on the properties of connectionist models and their relevance in areas other than language. As a stimulus to research and thinking, the chapter was enormously successful. We know much more about many things as a consequence of research that it inspired.

The past tense debate came to be seen as addressing some of the most important issues in cognitive science. This research has been discussed extensively elsewhere (see, e.g., McClelland & Patterson, 2002; Pinker & Ullman, 2002, and the accompanying replies). Our goal is to place the RM model in a broader context than was possible at the time it was published, making full use of the hindsight provided by the 25 years of research it stimulated. Our main point is that the chapter did not generate an intellectual firestorm because of what it claimed about the past tense, a genuinely minor aspect of the grammar of one language. Rather, it was the implications of the research, many of which were not explicitly stated in the chapter but were elaborated in later work, that created so much excitement and consternation. Our goal is to spell out those implications and examine the impact of the ideas the model embodied in areas that extend well beyond the past tense. The RM model and the reactions to it represent a very interesting episode in the history of ideas, a contemporary clash of paradigms in the study of language.[1]

2. The importance of the past

Most of the RM chapter was devoted to detailed descriptions of the model's architecture (a two-layer network trained using the delta rule) and behavior (generating past tense forms). Several properties of the model were noteworthy.

1. Whereas previous theories distinguished between rule-governed forms (known ones such as *bake-baked*, novel ones such as *dake-daked*) and "strong" forms (e.g., *take-took*), the RM model did not. The model took a phonological representation of a present tense verb as input and produced its past tense as output, the same procedure applying whether the input corresponded to a regular, irregular, or novel verb. This approach was quite surprising. A rule could generate *baked* and *daked* but not *took*. Rote memorization could handle known forms (*baked* and *took*) but not novel ones. The simultaneous demands of producing irregular past tenses (which must be learned) and generating novel forms (which could not have been learned) seemed to logically demand two distinct mechanisms. Rules, such as the one for the past tense in English, are a part of grammar. Forms that cannot be generated by rule are stored in the mental lexicon (Pinker, 1991). This rules-plus-exceptions view has great intuitive appeal; see, for example, Marshall and Newcombe (1973) and Coltheart, Davelaar, Jonasson, and Besner (1977) for similar arguments in a very different domain, pronouncing written English. Some of the appeal derives from the fact that it is guaranteed to cover all cases. The rules handle as many cases as the rules handle, and a backup system handles everything else. RM's "uniform procedure" was a radically different approach—no rule, no backup—that was intuitive only within the context of the PDP framework.

2. The model provided a novel account of generalization. Given a present tense as input, it produced the past tense as output. Generalization was merely the case in which the model generated output for a pattern on which it had not been trained (e.g., *nake*). This idea was remarkable because previously it was taken as a truth, nearly universally acknowledged, that generalization is accomplished by, indeed provides a strong evidence for, rules—a particular type of knowledge representation essential to language (Pinker, 1991, 1994, 1997, 1999).

3. The model suggested that this knowledge could be acquired through exposure to examples, via a weight-adjusting learning procedure. This aspect contrasted with theories in which language learning involves generating and testing hypotheses about rules, parameters, and other components of universal grammar (Wexler, 1998).

4. The model was constructed out of elements that were not specific to the past tense or to language. The past tense model, an early effort, was idiosyncratic in its details compared to the many PDP models that followed. Still, it consisted of generic elements (e.g., units activated via weighted connections, adjusted on the basis of feedback) thought to capture some important properties of knowledge representation, learning, and processing modeled on how the brain works. Such models can then be applied to different domains (e.g., language, vision) and tasks (e.g., generating past tenses, recognizing objects). The language-specific elements in the RM model were the input and output representations (phonology) and the task (learning and computing the English past tense). The model thus instantiated the idea that language, like other aspects of cognition, involves recruiting general mechanisms in the service of particular tasks. This approach contrasted with the standard view that language exhibits unique properties and thus cannot be the expression of more general cognitive and learning capacities (Chomsky, 1985).

5. The model focused on performance (it performed a task) rather than on characterizing competence (the grammar of the past tense). Chomsky had long asserted that characterizing the nature of language ("competence") was the logical prerequisite for investigating how it is learned, used, and represented in the brain ("performance"). Chomsky considered this argument for the epistemological precedence of grammar to be utterly obvious, and it has been enormously influential. For many years, linguistic theorizing that focused on characterizing competence grammar held the high ground in the study of language. Psycholinguistics became the study of how grammar is used in comprehension and production; acquisition, the study of how grammar is acquired; and neurolinguistics, the study of the brain bases of grammar. On this view, what we know when we know a language is logically distinct from how the knowledge is acquired and used. This view was especially prominent in the era when the PDP books appeared (e.g., Lightfoot, 1982), and it continues to underlie much research (e.g., Guasti, 2002; Prince & Smolensky, 2004).

The past tense model reversed this relation: The goal was to understand language in terms of the mechanisms that support its acquisition and use, and their brain bases. Competence theories are generalizations about properties of some of the output of this system (sentences, primarily). This level of theorizing is useful for some purposes; for example, it facilitates the description and comparison of languages. Rumelhart and McClelland rejected the further claim that grammar provided the necessary starting point for attempts to understand the neural and computational mechanisms that support language acquisition and use. The RM model launched what came to be called an "emergentist" view of language and cognition (MacWhinney, 1999; McClelland et al., 2010), in which capacities such as language emerge from the simpler, noncognitive operations that govern learning and processing in the brain. What we know when we know a language is the outcome of a developmental process, shaped by biology, experience, and goals (i.e., the communicative functions of language).

6. The modeling framework emphasized the relevance of neurobiology to understanding cognitive processing in general, and language in particular. The competence approach attempted to identify the essential characteristics of language through the analysis of primary linguistic data, mainly the grammaticality judgments of experts. The successful theory would rationalize the behavior of an idealized speaker–hearer, providing a basis for investigating language development, use, and brain bases. This research program, which originated in an earlier era, treated the cognitive and neural systems that give rise to language as a black box. The approach gained support from the functionalist argument that the mind's software could be characterized independently of the machine it happened to run on. At the same time, it was also commonly asserted that the essential properties of the human machine could be deduced from a proper theory of the structure of language, in advance of direct inspection of the neurobiology; see, for example, Bickerton's (1984) bioprogram hypothesis or Lightfoot (1982). From behavioral studies of the past tense, for example, we could be sure the brain somehow instantiates rules and an associative memory system. Thus, the research program involved working from competence to

performance, behavior to brain, outside to in. Properties of the brain could be "reverse-engineered" from properties of grammar.

The RM model was a first step toward characterizing the contents of the box, with language (exemplified by the past tense) as its output. The PDP approach disavowed the hardware–software dichotomy, holding that the biological and computational mechanisms that underlie the acquisition and use of language and other human capacities were critical to understanding their character. These mechanisms could begin to be investigated directly, rather than reverse-engineered via linguistic analysis.[2] Neurobiology could be linked to behavior via computational models of basic mechanisms of learning, processing, and knowledge representation. Research could proceed at all three levels simultaneously, with findings at each level constraining hypotheses about the others. This was perhaps the most radical aspect of the RM model. It challenged assumptions about how to go about understanding language that were the foundation of modern research on the topic.

7. Finally, the RM project was important because it illustrated the value of a particular type of modeling methodology. The goal of much research in psychology, and particularly in the study of language, is to develop high-level theories that provide insight about essential characteristics of behavior and the factors that cause it to be the way it is. In practice, theories often function to provide generalizations that systematize a body of empirical phenomena. In the case of the past tense, a broad range of data were marshaled in support of the conclusion that knowledge of a language includes rules (such as the one governing the past tense) and a lexicon (the repository for the exceptions) as essential components. Only a theory that incorporated this distinction could explain facts ranging from the child's production of *brang* to the ill-formedness of *rats-eater* to brain-injured patients' errors in producing the past tense. Such theories are stated in words rather than spelled out in formal or computational detail, leaving open many questions about the underlying mechanisms that give rise to behavior, and how they develop and function given human biology and experience.

Rumelhart and McClelland took a different approach. The PDP books described a general theory about the representation, acquisition, and processing of information, derived from insights about brain and behavior. The elements of the PDP framework were general rather than language specific, and they were developed and assessed using computational models such as the one for the past tense. Such models require far more detailed commitments than do the informal theories. This level of detail is not a mere nicety; it serves several functions. First, it is required in order to determine if proposed mechanisms are capable of giving rise to behavior in the intended ways. A high-level theory might rely on theoretical distinctions that have intuitive appeal but cannot be implemented in a way that is consistent with empirical findings and with other facts about people's capacities and experience (e.g., the implementation might require computations that are beyond human capacities). Second, the implemented model might give rise to target behavior in ways that would not otherwise be apparent. For example, the idea that a single mechanism could give rise to both the rule-governed past tense and irregular past tenses was counterintuitive at the time it was proposed. Merely asserting that such an outcome could

be achieved within a system employing the PDP principles would have little didactic impact; an implemented model that instantiates the behavior is necessary. Third, a model provides a strong way to test a theory, via comparisons between the model's behavior and target empirical phenomena (concerning, e.g., the formation of the past tense, or breakdowns in its use due to brain injury or disease). In time, much of the debate about the past tense came to focus on the validity and utility of the simulation modeling methodology, particularly the PDP variety. These concerns diverted attention away from substantive claims about the past tense, about language, and about the major theoretical tenets of the PDP framework.

In summary, the past tense model was important because it provided the vehicle for challenging two classes of assumptions: ones about how language should be studied (e.g., grammar leads performance; verbal descriptions vs. mechanistic computational models), and ones about the nature of language itself and how it is acquired and used.

3. The problem with the past

All of these claims were vigorously contested. In brief:

3.1. Regulars and irregulars are different

Studies of English and other languages suggested that rule-governed forms and irregular forms differ at every turn, indicating that they are distinct types of knowledge governed by different principles (Pinker, 1991). The cases that deviate from the past tense rule are handled by a second module, the lexicon. This two-system account built on a long-standing assumption in linguistic theory that knowledge of language consists of grammar plus a lexicon that is the repository for language-specific idiosyncracies, such as irregular morphology (Halle, 1973). Pinker and Prince (1988) assumed this standard notion of the lexicon, "with a twist," namely that it was organized as an "associative net" rather than a dictionary; this account underwent further developments summarized in Pinker (1999).

The central claim of this approach is that the two modules encode different types of information, governed by different principles and realized in different neural substrates, predicting that they will behave differently. As an example, past tenses differ in frequency of use. Although frequency has a large effect on the production of irregular past tenses, it has little or no effect on regulars (Prasada & Pinker, 1993). Irregulars show the effect because they are stored in the associatively structured lexicon. Regulars do not show the effect because they are generated by a rule that does not access word-specific information such as frequency. The 1990s saw the accumulation of an extensive body of evidence from several languages concerning multiple phenomena taken as evidence that rule-governed forms and irregulars are fundamentally different (see Pinker, 1999, for summary), contradicting RM's use of a "uniform procedure."

3.2. The account of generalization failed

In the RM model, generalization depended on phonological properties of the patterns on which it had been trained. People, however, can easily assign past tenses to highly irregular patterns, for example *ploamph-ploamphed*. The past tense rule was said to be blind to phonological properties of the verb to which is applied, whereas the RM model definitely was not. This inappropriate sensitivity to phonology was seen as contributing to anomalies such as producing *tour-toureder*.

3.3. The account of acquisition failed

Facts about children's acquisition of the past tense were taken as inconsistent with the model's basic assumptions. For example, Marcus et al. (1992) found that whereas children initially memorized a small vocabulary of regular and irregular verbs based on frequency of exposure, mastery of the past tense rule and production of overregularizations were unrelated to properties of caregiver input. The rote learning of past tenses, the reliable production of regular past tenses coincident with the appearance of overregularizations (e.g. *run-runned*), and the eventual unlearning of the overregularizations were taken as a strong evidence for two types of knowledge acquired on different schedules by different principles (Marcus et al., 1992). RM's own account of this "U-shaped" developmental sequence was inadequate because it relied on changing the training regime in ways that did not reflect children's experience.

Many other arguments against the RM approach to language acquisition unrelated to the past tense also appeared. The most important were based on evidence that children had control over various aspects of grammar as young as they could be tested (Chien & Wexler, 1990) suggesting they were known innately and arguments that language exhibits properties that cannot be learned from experience because there is literally no evidence for them in the input (Crain, 1991). Formalizations of the "logical problem of language acquisition" yielded results indicating that languages could not be acquired unless core properties were innately specified (Baker, 1981; Gold, 1967).

3.4. Language is special

The RM model was built out of generic parts; the model could as well have been configured to perform a nonlinguistic task. This property contradicted the mainstream view that language has unique properties and thus cannot be a product of general capacities to think and learn (Piattelli-Palmarini, 1994). The model's treatment of even a relatively simple aspect of one language using general rather than domain-specific mechanisms was a nonstarter for people for whom the uniqueness of language was an indisputable fact.

3.5. Implementational limitations vitiated substantive claims

Every simulation model is limited in scope: No model addresses all aspects of any phenomena. The basic methodology is to simplify some elements of the implemented

model in order to be able to investigate other issues in depth. The costs associated with these necessary simplifications are thought to be outweighed by the benefits of implementing an explicit model discussed above. For example, the RM model used a simplified representation of phonology in order to investigate more basic issues about the nature of the past tense. These built-in limitations create a variety of concerns. First, they sanction the easy criticism that the model cannot handle phenomena that are real but of no immediate relevance to the past tense (e.g., the RM model's "Wickelphonology" did not distinguish *papatap* from *patapap*; Pinker & Prince, 1988). Second, every model is literally false at some level of detail. It may capture core phenomena that were the immediate focus but fail to capture related phenomena. The failures at the edges of the model's performance call into question the validity of the account of the main phenomena of interest. The extreme version of this view is that various theoretically uninteresting details of the implementation are responsible for a model's seeming successes, as Lachter and Bever (1988) argued regarding the RM model. Third, a model might succeed only because its scope is limited, that is, exactly because it addresses a narrow range of phenomena. A model of categorization might "work" if it only has to distinguish between living and nonliving things but fail if a broader range of cases are considered. A model of the past tense might account for the phenomena to which it was applied but fail when additional phenomena are considered. Of course, the same kinds of concerns arise in connection with other types of theories, as we will see.

These concerns called into question not merely RM's treatment of the past tense but the value of connectionist computational modeling in general (McCloskey, 1991; McCloskey & Cohen, 1989; Ratcliff, 1990; see also Seidenberg, 1993). Alternatives to Rumelhart and McClelland's account of the past tense rarely took the form of models to which theirs could be compared (although see Hahn & Nakisa, 2000, and Nakisa, Plunkett, & Hahn, 2001, for exceptions). The use of a more informal, noncomputational style of theorizing (as in Pinker, 1999) obviates the limitations of the simulation approach but retains the limitations that modeling was intended to address.

3.6. Commitment to competence theories

The competence assumption remains deeply embedded in the study of language and cognition. Chomsky's assertion that "there seems to be little reason to question the traditional view that investigation of performance will proceed only so far as understanding of underlying competence permits" (Chomsky, 1965, p. 10) achieved the status of truism within mainstream linguistics and its offshoots. It took outsiders like Rumelhart and McClelland to challenge this view. The challenge was turned back by treating the RM model as though it represented the upper limits of the general framework rather than demonstrating some aspects of it. The model's failures could then be taken as evidence against Rumelhart and McClelland's inversion of competence and performance. The competence approach gained further support from Marr's (1982) proposal regarding levels of theorizing, particularly the computational level. Like Chomsky's competence theory, the computational level also characterized the nature of a problem (e.g., in vision or language) independent of the

procedures used in solving it or the machinery that runs them. McClelland et al. (2010) discuss the similarities between these ideas in the context of recent cognitive Bayesian models. The approaches achieve a natural convergence in recent accounts of language acquisition in which competence models of grammar are embedded within a Bayesian computational-level theory of the task (e.g., Perfors, Tenenbaum, & Regier, 2011).

The general conclusions that many drew from this body of work were that (a) a range of facts spoke against basic assumptions of the RM model because (b) the model instantiated the wrong approach to the past tense and to language in general, due to the fact that (c) it inherited the limitations of earlier associationist approaches to which Chomskyan linguistics had been a correction. Thus, (d) the flaws in the RM model reflected intrinsic problems with the approach and so could not be fixed. Attempts to develop new models to address the limitations of the RM model were pointless because the framework on which all such models were based was fatally flawed.

4. Where the past went

Many studies of many phenomena in many languages have addressed the validity of these claims and counterclaims over the past 25 years. This work seems to exemplify how modern research on language should proceed: It is theoretically driven and incorporates multiple types of data drawn from varied languages, components of languages, tasks, and subject populations. It attempts to directly compare competing accounts that make different assumptions. Many models of the past tense and related phenomena were developed, addressing limitations of the RM model and advancing their approach. The rule-based theory presented in Pinker and Prince (1988) was developed further as well (Pinker, 1991, 1999). All of this effort has unquestionably yielded new knowledge and insights. Yet the fundamental questions about the structure and use of language which the RM model brought out so effectively cannot be said to have been resolved. Even close analyses of circumscribed phenomena for which the different approaches made very different claims failed to yield a Popperian knockout. Examples that illustrate this point are discussed below. Despite extensive research, the science has not converged on consensus. Why not? We think the answer can be found by examining what the debate was about.

Was the debate literally about the past tense? The past tense has been called the "Drosophila of language research" (Pinker, 1999): It would act as a "model system," the extended study of which would yield essential insights about language. The Drosophila analogy is ill fitting, however.[3] Although much has been learned about genetics and other aspects of biology from studying Drosophila, what can be concluded about language from studies of the English past tense is severely limited. Drosophila are interesting precisely because of the characteristics they share with a large part of the biological world. In contrast, the past tense is notably lacking the structural complexity, prominent in syntax and phonology, that led to positing highly abstract underlying structures, including ones thought to be unlearnable (i.e., innate) because there is no overt evidence for them. The English past tense is atypical even as an example of inflectional morphology. The system

is so simple that it can indeed be described by a small number of rules (with exceptions), but this treatment scales up poorly to the more complex morphological systems found in many languages (e.g., Mirković, Seidenberg, & Joanisse, 2011). Thus, the problem with taking the past tense as a model system is that it might obscure broader generalizations that can only be identified by addressing other aspects of English, other inflectional systems, and other properties of language.[4]

Was the debate about rules? Both RM and Pinker focused on the status of rules, but rules have not been an essential ingredient in linguistic theorizing since Chomsky (1981). For example, influential later theories incorporated systems of "violable" well-formedness constraints (Prince & Smolensky, 2004), which exhibit both greater formal similarity to and historical continuity with the PDP approach than the traditional concept of "rule." For many linguists, the debate over whether the past tense was rule governed or not was largely irrelevant to their theoretical interests.

Was the debate about nativist versus empiricist views of language? It is sometimes posed as such (e.g., Alegre & Gordon, 1996; Ramscar & Dye, 2010), but in reality both approaches contained elements of each. The basic processing and learning mechanisms of the RM and other PDP models cannot themselves be learned. If such models were truly tabula rasa, they obviously could not function; a blank slate would also be an inert slate. Conversely, inflectional morphology (including the past tense) is a poor candidate for innate grammatical knowledge. Inflectional systems are not universal; many languages lack them entirely. Whether the language to which a child is exposed makes use of inflectional morphology, for what linguistic purposes, and how it is realized are properties that must be learned. Whereas the English inflectional system lacks the highly abstract, putatively unlearnable properties of language, it is very similar to an invented system that is learned via instruction, English orthography. This commonality allowed Harm and Seidenberg (2004) to use the same network to generate both the pronunciations of words from print and the past tenses of verbs. Thus, the past tense has little bearing on the question of innate linguistic knowledge.

The important debate was actually about the validity of two different conceptions of the nature of language, one treating it as distinct from the rest of cognition, governed by domain-specific principles, the other treating it as continuous with the rest of cognition, governed by principles that apply broadly to brain and behavior. However, the debate's importance extended beyond the past tense, beyond the status of rules or details of the RM model, and beyond language. It was also a debate about what there is to explain about complex phenomena such as language. The past tense may not exhibit all of the myriad complexities of language, but it is highly representative in one critical respect: Past tense formation is systematic but admits exceptions that deviate from the main pattern to varying degrees. Rumelhart and McClelland (1986) took note of this property, which Seidenberg and McClelland (1989) termed "quasiregularity." The critical property is not merely that rules have exceptions, but that *the exceptions are not arbitrary*: They share structure with the rule-governed cases and with each other (Daugherty & Seidenberg, 1994). For example, the past tense of the verb *make* is irregular. However, it is *made* (/meId/), not *glorp*. *Make* and *made* are related: They have the same onset and vowel, as do the regular *fake-faked*.

Moreover, if there were a verb "to may", /meId/ would be its regular past tense; thus, *made*, like many regulars, ends in the coda /d/ that is one of the realizations of the regular past tense inflection. Processes underlying the emergence of forms such as *made* were simulated by Lupyan and McClelland (2003). McClelland and Patterson (2002) characterized the phonological overlap between regular and irregular forms in detail (see also Halle & Mohanan, 1985, who used rules to generate past tenses such as *made* in order to capture the overlap with the present tense). Many other aspects of language have been argued to have a quasiregular character: phonology, derivational morphology, and grammaticality itself (Bybee & McClelland, 2005; Seidenberg & Gonnerman, 2000; Allen & Seidenberg, 1999).

There are two ways to approach such observations. One is to take the fact that regulars and irregulars share structure as prima facie evidence that they are part of the same quasiregular system, governed by the same principles of learning and processing. The other is that "the exceptions prove the rule." That is, rules are an essential element of language, the past tense is obviously rule governed, and the exceptions merely show that our brains are sufficiently flexible and capacious to accommodate irregularities that arose for assorted historical and functional reasons (e.g., reducing *hitted* to *hit* under pressure to increase ease of articulation). Rumelhart and McClelland, cognizant of work by Bybee and Slobin (1982), took the first approach, using the PDP principles. Pinker and colleagues took the second approach, building on a long-standing distinction between grammar and lexicon ("The lexicon is really an appendix of the grammar, a list of basic irregularities." Bloomfield, 1933, p. 274). Is the core property of language its systematicity (Fodor & Pylyshyn, 1988) or quasiregularity (Seidenberg & McClelland, 1989)?

Whatever Rumelhart and McClelland's original goals, they were rapidly overtaken by dissections of the model's limitations. Mere proof of concept did not carry the day. The RM model and Pinker and Prince's (1988) reaction to it succeeded in stimulating years of research. On the empirical side, the central question was whether behavior and (later) brain respect the categorical distinction between regular and irregular forms asserted by the two-module account, or whether the quasiregular characterization is more accurate. On the theoretical side, attention focused on determining whether the limitations of the RM model were related to its specific goals and implementational details, or whether they were "irremediable" because they instantiated a bankrupt "connectionist ideology" (Pinker & Prince, 1988, p. 82). We summarize representative findings in these areas below, and then consider the broader impacts of the RM model as seen in contemporary research.

5. What are the facts?

The field awaits an inclusive, nonpartisan, analytic review of the many studies stimulated by the RM chapter, for which this article is not the appropriate vehicle. As a rule (so to speak), empirical phenomena do not exhibit categorical differences between rule-governed forms and exceptions predicted by the two-module theories. The most common pattern, without question, is that regulars and irregulars are both affected by a given property or condition, but not equally. The predictions from the rule-based theory are stated

as categorical (e.g., certain forms "never" occur, a stimulus property has "no" impact on use of the rule, patients are "selectively" impaired in the use of one of the two components, and so on). The plausibility of these predictions is established by a clear example or two. The experiments involve assessing whether the same pattern is maintained over many such examples. The actual data are different: Prohibited forms occur but with lower frequency or lower-rated acceptability; the stimulus property has a smaller impact on one type of past tense but not no effect; there are intermediate cases that fall between the regular and irregular extremes; patients are impaired on both types of forms but more impaired on one than the other. Thus, it would be more accurate to say that regulars and irregulars dissociate, not "from virtually every point of view" (Pinker, 1991, p. 532), but from most views, somewhat. There are a few case studies of brain-injured patients who most closely adhere to a categorical pattern of dissociation (e.g., patient FCL: Ullman et al., 1997; Miozzo, 2003), but the interpretation of these cases is highly controversial (see Bird, Lambon Ralph, Seidenberg, McClelland, & Patterson, 2003; Joanisse & Seidenberg, 1999; Woollams, Lambon Ralph, Plaut, & Patterson, 2007). Such cases notwithstanding, the striking fact is how difficult it is to find examples of the types of dissociations predicted by the two-module theory.

Everyone knows the data pattern this way. What differs is how this pattern has been addressed. For the RM approach, the absence of such dissociations is expected because all forms are generated by a common system. Graded effects arising from interactions among multiple soft constraints are the bread and butter of PDP networks. The research questions then concern the exact structure of this system, how it is learned, the factors that influence its functioning, and how they play out with respect to verbs along the quasiregularity spectrum.

On the other side, failures to observe strong dissociations might seem injurious to a theory in which the central claim is about independent modules with wholly different properties. They have not proved fatal, however. For proponents of this theory, deviations from categorical predictions merely show that performance is affected by other factors *unrelated to their core theory*. If the deviations are unsystematic, they can be attributed to measurement error. If the deviations are systematic, they may be due to additional factors arising from outside the linguistic systems in question. It does no harm to the theory to identify such additional factors, on this view; to the contrary, it yields a fuller accounting of the data.

Indeed, such an account might prove to be correct. The danger, of course, is that the theory might be protected from disconfirmation by an extra layer of explanatory insulation. Data can always be said to fit a theory if there are additional mechanisms to deal with the cases for which the main predictions fail. It then becomes important to examine how well the data fit the original predictions, whether the deviations from predicted patterns are systematic, and whether some other theory could account for the full range of cases without special pleading.

Here is a prototypical example. Linguistic rules are said to have specific properties that distinguish them from other types of knowledge. Grammatical rules like the one for the past tense are triggered by grammatical conditions, in this case a grammatical marker for

the past attached to the base form of a verb. Because the triggers are grammatical, the rule does not have access to the phonological and semantic properties of the verbs to which they apply. In contrast, the lexical module, which encodes phonological and semantic relations among words, is responsible for irregular past tenses. These properties can then affect production of irregular past tenses, but not regulars. Those are clear predictions; what are the data?

Consider phonology first. Novel verbs such as *plip* and *ploamph* (Prasada & Pinker, 1993) differ greatly in the extent to which they resemble other words. *Plip* overlaps with close neighbors such as *slip* and *plop*; *ploamph* has no close neighbors. What are the past tenses of these novel verbs? *Plipped* and *ploamphed*, presumably. This observation supports the claim that rule application is independent the phonological properties of verbs. Prasada and Pinker (1993) conducted experiments that confirmed this pattern with a larger number of items. The past tense rule's apparent insensitivity to phonology contrasted with the RM model, for which phonological properties of the verb greatly affected performance, with poorer performance on phonologically atypical items, especially nonwords.

These findings establish that generation of the regular past tense is sometimes unaffected by phonological properties of the present tense verb. Are there cases in which phonology does affect rule-application? Of course. Knowledge of such cases predates the modern past tense debate. In the ur-study in this area, Berko (1958) examined the generation of past tenses for novel verbs such as *bing* and *gling*, which are phonologically similar to a cluster of irregular verbs that includes *sing*, ring and *bring*. If rule application is blind to the phonological properties of the stem, the past tenses should be *binged* and *glinged*. However, adult subjects produced *bang* or *bung* as the past tense of *bing* (50% of responses) and *glang* or *glung* as the past tense of *gling* (75% of responses). Bybee and Moder (1983) identified the phonological characteristics of several subgroups of irregular past tenses that govern the exact realizations of such "irregularized" past tenses (see also Bybee & Slobin, 1982). Thus, rule application is both insensitive to phonology, in cases such as *ploamph*, and sensitive to phonology, in cases such as *gling*.

The two-module theory retained the idea that rule application is phonologically blind by invoking an additional factor: sometimes the perceptual similarity of a novel verb to known irregulars interferes with the normal process of rule application, producing irregularizations such as *glang* (Pinker, 1999). An alternative approach would be to seek a covering generalization that subsumes the full range of cases without complication. The broader generalization concerns the phonological properties of verbs and the existence of subpatterns in the mappings between present and past tenses. The existence of such subpatterns among the irregular pasts is well-known. However, using an agorithmic rule-discovery procedure, Albright and Hayes (2003) found such patterns, which they termed "islands of reliability," among regular verbs as well. These too affected generation of past tenses for novel verbs and well-formedness judgments.

These findings (see also Albright & Hayes, 2006) are important. First, they show that rule-application is not "blind" to the phonological properties of regular as well as irregular verbs. Second, they undermine the claim that presence versus absence of sensitivity to phonological similarity shows that irregulars and regulars are governed by distinct subsystems.

Such results might be handled within the two-module theory by assuming that the "associative net" represents phonological similarities among all verbs (indeed, all words), and that it is used in producing all past tense forms ("regular," "irregular" and novel). At this point, however, the theory collapses into Rumelhart and McClelland's original proposal. What, then, is role of the past tense rule? The phonological contingencies governing past tense generation are complex, hardly visible to the naked eye; it took close studies such as Bybee and Moder (1983) and Albright and Hayes (2003) to discover many of them. PDP networks are systems for learning and representing the statistical structure underlying the past tense and many other types of knowledge. This extensive set of contingencies, encoded by a large number of weighted connections between units, is difficult to describe concisely. The classic past tense rule serves the heuristic function of providing a simple verbal description of a major pattern within this system. It is not literally correct, but it can be communicated more easily than the structure inherent in a PDP network and it can be useful for some purposes (e.g., cross-linguistic, typological comparisons of inflectional systems). It serves a descriptive function similar to the first principal component in a factor analysis.

To summarize, consideration of a broader range of empirical phenomena, many discovered as a result of studies stimulated by the two-module theory, showed that the role of phonological similarity in past tense generation did not provide a basis for distinguishing rule and lexicon modules.

Now consider phonology's complement, semantics. Rule application also lacks access to a verb's meaning. This is thought to be a good thing because semantics is not a reliable predictor of the past tense. Verbs exhibit highly diverse semantics. The vast majority have regular past tenses; semantic factors do not predict which verbs have irregular past tenses, or their exact forms. Pinker and Prince (1988) surmised that a system that predicted past tenses from semantics would be befuddled by verbs such as *slap*, *strike*, and *hit*, which are similar in meaning but form the past tense in three distinct ways (regular, vowel change irregular, no change irregular, respectively). Moreover, people can easily generate past tenses for nonce forms, which lack semantics (e.g., *ploamph*). These observations support the claim that rule application is independent of the semantic properties of verbs.

Again, granting these observations, are there any cases where semantics does matter? Of course. One clever example: Ramscar (2002) investigated the generation of past tenses for novel forms such as *sprink*, which were introduced in story contexts. The question was whether subjects would prefer *sprinked* or *sprank*. The content of one story related to eating and drinking and was intended to implicitly prime the word *drink* (irregular past: *drank*) which was not explicitly mentioned. The content of the other story related to eye problems and was intended to prime the word *blink* (regular past: *blinked*), not explicitly mentioned either. In the drinking context, subjects preferred *sprank*; in the eye context, they preferred *sprinked*. This result replicated with a second novel verb *frink*. Thus, the semantics of the primed verb influenced the generation of novel past tenses.

Is Ramscar's result necessarily fatal for the rule-based theory? If one's prior probability for the two-module theory is high, the likelihood of the evidence must be low, as can be assured by faulting minor aspects of the procedure and results (Huang & Pinker, 2010), yielding little change in the theory's posterior probability. It is only one experiment. The number of nonce verbs was small. Perhaps demand characteristics of the experiment caused subjects to think the experimenter wanted them to use semantic information in generating responses. Such concerns could be addressed with additional experiments (e.g., the other ones in Ramscar, 2002; also Ramscar, Dye, & Hübner, 2013, who replicated these findings using on-line reading tasks rather than preference judgments).

Regardless of the results, one could again appeal to contamination: Rule application really is blind to semantics (and phonology) but people are fallible processors and so nonwords can sometimes be sucked into the associative net, trumping the rule but not the theory.

In summary, the strong claims about categorical differences between rules and lexicon are not literally empirically true, as shown in numerous experiments. The theory can be saved by attributing the nonconforming data to complications separate from the core theory. This approach to data handling is easily recognizable as an extension of Chomsky's competence-performance distinction. Characterizing linguistic competence was thought to be a prerequisite for investigating how this knowledge is acquired, used, and represented in the brain. However, all of the data about competence grammar arise from performance, principally expert judgments of the well-formedness of sentences. This created a paradox: Performance data were needed in order to uncover grammar, but grammar was needed in order to understand performance—including how people make grammaticality judgments and why such judgments often conflict (Schütze, 1996). In practice, grammatical theories made strong empirical claims, supported by clear cases. Deviations from predicted patterns were attributed to performance factors unrelated to grammar. The vagueness of the boundary between competence and performance provided considerable insulation from disconfirming data.

Whereas theoretical linguists were mainly interested in performance data as a tool for uncovering properties of grammar, Pinker's innovation was to extend their approach to the study of performance itself. Whereas Chomsky distinguished between knowledge (competence) and use (performance), Pinker and colleagues drew a further distinction between the aspects of performance that are relevant to their performance theory and ones that are not. The result was a *competence theory of performance*. The two-module theory does not accurately characterize people's performance but rather what performance would be like if only we were more reliable language users, immune to other factors.

The alternative approach, initiated by Rumelhart and McClelland, shared with Chomsky the goal of explaining how language is acquired, used, and represented in the brain, but rejected the distinction between knowledge and use of language. This distinction, which seemed so obvious and necessary to Chomsky, working with concepts such as grammars, parsers, and language acquisition devices, is not intrinsic to the PDP framework, in which the representation of knowledge is not distinct from the processing mechanisms that make use of it. Past tense models also ambitiously attempted to incorporate

aspects of performance that had to be treated separately in competence theories of performance.

6. The importance of semantic and phonological constraints

In the examples discussed above, semantic and phonological properties of verbs exerted effects where the two-module theory says they should not. Plausible post hoc accounts of these effects can be generated in each case, taken individually. The great value of the past tense debate is that there is so much data about so many of its properties that it is possible to evaluate whether there is a pattern to the deviations from expected results. Looking at these findings, it becomes apparent that semantics and phonology are implicated repeatedly. We briefly summarize two additional cases, concerning seemingly disparate phenomena, in order to convey the generality of this pattern. We also describe a third case, concerning impairments following brain injury, summarized by McClelland and Patterson (2002). These and other discoveries contributed to the development of the modern descendants of the Rumelhart and McClelland theory, in which the phonological processes they described are part of a larger lexical system that also computes meaning and, in literate individuals, orthography (Harm & Seidenberg, 2004; Joanisse & Seidenberg, 1999; Plaut, McClelland, Seidenberg, & Patterson, 1996). Lexical knowledge consists (in large part) of knowledge of the relations between different codes: semantics, phonology, orthography. The network that encodes these relations is used in performing a variety of tasks: comprehending the meaning of a word, generating pronunciations from print, generating the form of a word from meaning, and others, including generating the past tense. The network computes the best-fitting form, given the task and knowledge encoded by the weights. This constraint satisfaction process, for which semantics and phonology are the most salient factors, is a fuller realization of Rumelhart and McClelland's "uniform procedure."

This network is also part of a sequential processing system that tracks distributional properties of words (McClelland, St. John, & Taraban, 1989; Seidenberg & MacDonald, 1999) which provide information about word and utterance form and meaning. Critically, the core mechanisms of the constraint satisfaction approach subsume behaviors that the two-module theory attributes to outside factors.

7. Compounding

Whatever theory explains the past tense should simultaneously account for the formation of singular and plural nouns, the other major part of English inflectional morphology (Kiparsky, 1982). Singular nouns in English, like present-tense verbs, are not overtly inflected (e.g., *book*, *tie*); regular plurals follow a rule (e.g., *books, ties*); the inflection has a consistent spelling (*-s*), the phonological realization of which depends on the final phoneme of the singular noun; people readily generalize to novel forms (e.g., *nust-nusts*),

even phonologically weird ones (*ploamph-ploamphs*); and a minority of plurals are irregular (e.g., *man-men, ox-oxen, goose-geese*), which overlap with regulars and with their singular forms in varying degrees. The two-module theory again holds that the singulars and irregulars are stored in the lexicon, whereas the regular plural is generated by rule. The data of interest concern the occurrence of these forms in compounds such as *book club* or *man hours*, which consist of a modifier (e.g., *book*) followed by a head (e.g., *club*). The critical phenomenon is illustrated in (1):

(1) (a) *rat-eater*
(b) *mouse-eater*
(c) **rats-eater*
(d) *mice-eater*
(e) *rat/mouse/mice-eaters*

Intuition suggests that out of all these cases, only (c) is ill formed. Why? The pattern is particularly puzzling because an animal that eats rats (i.e., a rat-eater) eats many rats, not merely one, yet **rats-eater* is the disallowed form. The data can be explained with four assumptions (paraphrased from Pinker, 1994):

A1. Singular nouns and irregular plurals are listed in the lexicon.
A2. Regular plurals are generated by rule.
A3. Compounds are formed by rule.
A4. The compounding rule applies before the plural rule.

(A1) and (A2) are carried over from the analysis of verbs; (A3) is the standard approach to compounding and other aspects of derivational morphology; (A4) is a further assumption about the order in which morphological rules are applied (in the "level-ordering" theory of Kiparsky, 1982, and Siegel, 1979). The four assumptions yield exactly the pattern illustrated in (1). Pinker (1994, 1999) discusses this case in detail, as evidence for both linguistic rules and a lexicon with discrete representational levels. Consistent with this analysis, Gordon (1985) found that young children dispreferred expressions such as *rats-eater* but were OK with *mice-eater* and *rat-eater*, from which he concluded that the level-ordered application of rules was part of innate grammatical knowledge. Note that this analysis makes no reference to the phonological or semantic properties of words. The rule again only has access to grammatical information: whether the word is a noun and whether it is flagged as having an irregular plural. The semantic similarity of *rat* and *mouse* suggested that semantics is not a relevant factor.

These phenomena have since been investigated extensively (e.g., Berent & Pinker, 2007; Haskell, MacDonald, & Seidenberg, 2003; Ramscar & Dye, 2010; Seidenberg, MacDonald, & Haskell, 2007). As before, results deviate from what should occur given the rules-lexicon dichotomy; that theory is preserved by attributing the deviations to other factors; the other factors are semantic and phonological, resulting in graded rather than categorical well-formedness. We will illustrate the basic findings using the much-repeated

rat-eater case; see the cited references for details. In brief, there are two decisive findings: irregular plurals are less acceptable and regular plurals are more acceptable than predicted by the rules and exceptions theory.

The level-ordering theory holds that singular nouns and irregular plurals can occur as modifiers but regulars plurals cannot. Asked to rate their acceptability ("goodness"), subjects indeed prefer *rat-eater* and disprefer *rats-eater* (Haskell et al., 2003). However, *mice-eater* receives intermediate ratings: better than *rats* and worse than *rat* or *mouse*. Corpus analyses also indicate that irregular plurals are used as modifiers less often than expected, were they as acceptable as singulars (Haskell et al., 2003).

Conversely, regular plurals are more admissible than predicted. *Rats-eater* sounds terrible but *sports reporter, weapons inspector, parks department* and *awards ceremony* are fine (as Kiparsky had noted). Corpus analyses also indicate that regular plurals occur more often as modifiers than expected from the **rats-eater* example. Although the similarity of *rats* and *mice* seemed to demonstrate that semantics is irrelevant, Alegre and Gordon (1996) attempted to develop a semantic account of apparent exceptions such as *sports reporter*. Their analysis is not fully successful—the semantic properties they identified are not sufficient to accurately predict which forms are in or out—but it suggests that semantics is relevant, perhaps via several interacting semantic distinctions.

Similar but complementary results were reported by Ramscar and Dye (2010), who focused on another phenomenon for which level-ordering appeared to provide an elegant account. Alegre and Gordon (1996) examined the interpretation of NPs such as *red rat eater*, which is ambiguous between two interpretations: *[red rat] eater* (eater of red rats, the NP/N structure) and *red [rat eater]* (rat eater that is red, the ADJ/NP structure). According to the level-ordering theory, *rats eater* is ill formed, as discussed above. It follows that the phrase *red rats eater* can only be interpreted *[red rats] eater;* level-ordering blocks the generation of *red [rats eater]*. Children heard such phrases, the two interpretations were explained and they were asked which meaning was most likely to them. The *[red rats] eater* interpretation (NP/N) was preferred over *red [rats eater]*, (ADJ/NP), consistent with level-ordering.

Ramscar and Dye (2010) performed several experiments involving these structures, with children and adults. The main finding is that subjects also prefer the NP/N construction in expressions such as *red mice eater*, which Gordon and Allegre had not tested. Level ordering allows both interpretations because both irregular plurals such as *mice* and singulars such as *rat* are listed in the lexicon. Ramscar and Dye's results show that there is a preference for the NP/N structure that is unrelated to regularity of the plural or to the level-ordering theory.

Level-ordering is a moribund theory, living on only in debates about the past tense. The phenomena, however, are important. The alternative to level-ordering is again to start with relations between form (phonology) and meaning (semantics). People prefer singular modifiers—that is, modifiers that have singular phonology (in English, a noun with no overt inflection for number) and singular semantics. Haskell et al. (2003) discuss the origin of the preference for singular modifiers (it arises from the fact that modifiers,

including adjectives, are not marked for number in English), and how the preference can be learned from speech to children. *Rat-eater* is acceptable because it is both semantically and phonologically singular (++); *rats-eater* is poor, because it is neither semantically nor phonologically singular (−); *mice-eater* is phonologically singular but semantically plural, producing intermediate-level acceptability (+−).

This account correctly predicts two other intermediate cases. Pluralia tanta, such as *scissors*, are phonologically plural (*scissors* would be the well-formed plural of *scissor*) and semantically singular (+−).[5] *Wolves* is another mixed case. The plural is rule governed insofar as it has the regular inflection and the noun is plural; hence, it should be disallowed along with *rats* and *bears*. However, it differs from other regular plurals because the rule applies to a deformed stem; *calf-calves* and *leaf-leaves* are other examples. Although these cases are rare in English, they are common in more complex inflectional systems (Mirković et al., 2011). In these voicing-change (VC) plurals, the final unvoiced consonant of the singular noun is voiced in the plural. Note that this stem deformation is not obligatory; compare *loaf-loaves* with *oaf-oafs*. Are the VC plurals rule-governed or exceptions, the only alternatives admitted by the two-module theory? Are they allowed as modifiers in compounds or not? In fact, they are more acceptable than fully conforming regular plurals such as *rats* but less acceptable than irregulars such as *mice* (Haskell et al., 2003).

The compounding facts are more extensive than presented here and are still debated (Berent & Pinker, 2007, 2008; Ramscar & Dye, 2011; Seidenberg et al., 2007). The main point is that the competing analyses of these phenomena are like the others discussed throughout this article. The two-module theory makes strong predictions that are partly borne out by the data; the deviations are explained by additional factors, which turn out to be semantic and/or phonological. Whereas the two-module theory only appeals to semantics and phonology to explain the part of the data that contradict their theory, the alternative is to derive the entire range of cases from a "uniform procedure" in which phonological and semantic constraints are the primary determinants of well-formedness (Joanisse & Seidenberg, 1999).

8. Creating novel verbs

People use existing knowledge to create new words, an example of the productive, creative nature of language. Constraints on the formation of novel verbs were offered as another, independent source of evidence for the distinction between grammar and lexicon, and against the network approach (Kim, Pinker, Prince, & Prasada, 1991; Pinker, 1999). The argument is elegant and only slightly technical. Consider the word *fly*.

1. *Fly* is a verb, the primary meaning of which is, roughly, "to move through the air using wings." Its past tense is irregular, *flew*.
2. From this verb, a derived, "deverbal" verb was created. This form occurs in the idiom "to fly off the handle." Its past tense is also *flew*.

3. *Fly* is also a noun, referring to a baseball hit high into the air ("a fly to left field"). From this noun, a derived, "denominal" verb has also been created ("to fly out to left field"). What is its past tense?

In the two-module theory, a verb's past tense is formed by applying the rule unless it is marked as having an irregular past tense. Kim et al. (1991) investigated a further condition: the past tense assigned to a novel verb depends on its "derivational status," specifically (a) whether it is derived from a noun or verb, and (b) if derived from a verb, whether the source word has a regular or irregular past tense. If a source verb has an irregular past tense, the novel verb inherits this form. Thus, (2a) is grammatical but (2b) is not. If the source verb has a regular past tense, so does the derived form. The kicker concerns novel verbs derived from nouns. They cannot inherit the irregular past tense because they are not derived from the relevant source verb; hence, (2d) is allowed but (2c) is not. (The data are the average acceptability ratings for these items from their study ranging from low (1) to high (7), provided as a representative example).

(2)	(a) He flew off the handle.	6.8750
	(b) *He flied off the handle.	1.8125
	(c) *He flew out to left field.	3.9375
	(d) He flied out to left field.	4.2500

Kim and colleagues summarized these predictions as follows:

> The formal grammatical hypothesis predicts that only verbs with verbal roots in head position can have an irregular past tense form. All denominal verbs will have a regular past tense form, even if they are ultimately related to some verbal root, whereas all deverbal verbs with irregular past tense roots will have an irregular past tense form. (p. 184)

Thus, the title of their paper, "Why no mere mortal has ever flown out to center field."

These phenomena are important because they provide a simple example of the "structure dependence" of linguistic rules. The form of the novel past tense is said to be determined by an aspect of linguistic structure, the grammatical category of the source word. Such structure-dependent operations are thought to be an essential characteristic of language (Chomsky, 1981); networks such as RM cannot capture such effects, it is argued, because they do not represent the relevant symbolic structures over which the rules apply. However, the appeal to the derivational histories of words raises questions about whether people who have not studied the topic possess this knowledge or how it could be learned.

What are the data? Kim et al. (1991) recorded subjects' ratings of the acceptability of regular and irregular verbs in sentences similar to (2a–d). Overall means supported their predictions. There was a significant interaction between source of a novel verb (noun or irregular verb) and preferred past tense (regular or irregular). If the source was a verb with an irregular past tense, the irregular past was preferred for the derived form. If the source was a noun, the regular past was preferred for the derived form. However,

subjects' preferences varied considerably. For some verbs the predicted form was rated more highly; for some verbs, the predicted form was rated less highly; for some the ratings did not differ reliably (see also Daugherty, MacDonald, Petersen, & Seidenberg, 1993; Harris, 1992). People clearly dislike "flied off the handle" (2b above); they also prefer (2d) to (2c), but the theory says (2c) should be flagrantly ill formed rather than slightly less preferred. As a corpus count quickly confirms, "flew out to center field" may be disallowed by the theory, but people say it frequently.

An alternative hypothesis (Lakoff, 1987) concerns—you guessed it—the meanings of derived verbs rather than grammatical category of the source. The general idea is that an irregular past tense can only be extended to novel meanings that overlap with the existing meaning. The *fly* examples are confusing because *fly* has multiple senses whose meanings overlap in varying degrees. The treatment of meanings as discrete, static data structures derived from one another is itself problematic, as *fly* illustrates. Nor can their derivational histories be independently established in most cases. A clearer example is (3), another of Kim et al.'s (1991) cases. Here the homophones are *steal* (verb) and *steel* (noun), the semantics of which are clearly unrelated.

(3) (a) John stole second base.
(b) *John stealed second base.
(c) *John stole himself for the expected criticism.
(d) John steeled himself for the expected criticism.

The novel deverbal form is (3a), stealing a base (analogous to the earlier flying off the handle). The denominal form is (3d). The main finding is simply that *stole* makes a very poor past tense for the verb "to steel." For Kim and colleagues, this follows from the fact that *steel* (the verb) derives from a noun. *Stole* can be used for stealing a base because it derives from a verb. Under the semantic account, (3c) is infelicitous because *stole*'s meaning is unrelated to *steel*–noun, verb or otherwise. In contrast, *stole* is acceptable in the baseball sense (3a) because it overlaps with the original meaning of *steal*, far more than it overlaps with *steel*: taking a base you haven't "earned" is a form of theft. It is not that a word form such as *stole* cannot acquire additional semantically unrelated meanings (viz., mink stole). Rather, it is infelicitous as a verb with a wholly unrelated meaning because it cannot be severed from its association with *steal*.

In summary, phonology and semantics again emerge as the major constraints on the target phenomena. Derivational status is confounded with semantic similarity: Other factors aside, a verb derived from an existing verb will usually be closer in meaning to the original verb than will a verb derived from a semantically unrelated noun.

Kim et al. (1991) were aware of the semantic distance hypothesis. To assess this hypothesis, they obtained ratings of the similarity between the meaning of each novel verb stimulus and the "central" meaning of an existing verb or noun. In a multiple regression, this distance measure did correlate with acceptability judgments; however, grammatical category of the source word accounted for significant additional variance. From their perspective, this result indicates that, when the semantic confound is factored out,

derivational history is revealed as the determining factor. As before, the alternative is to consider whether semantic and phonological factors can explain the entire range of results. This possibility was taken up by Harris (1992), Daugherty et al. (1993), and Ramscar (2002; Ramscar et al., 2013). Three important issues should be mentioned:

1. Regression analyses are only as good as the factors used as predictors. Kim et al. (1991) limited their test of the semantic distance hypothesis to the distance between a novel meaning and a single meaning intuited to be the "central" meaning of the source verb, attributing the residual variance to the only other factor considered, grammatical category. Other researchers assessed relations between novel and existing meanings in more depth, examining, for example, distance to any existing verb meaning, and distance between a denominal verb and its source noun. Often the derived meaning was closer to a different meaning than to Kim et al.'s designated "central" meaning. Reviewing these findings, Ramscar (2002) concluded that derivational status is relevant only insofar as it is confounded with semantic and pragmatic factors governing the creation of novel verbs.

2. Kim et al. found additional support for their theory using nonwords as novel nouns and verbs. Nonwords seem to rule out semantics, because they do not have established meanings. However, these nonce words were introduced in meaningful linguistic contexts, used to establish whether the source nonword was a noun or verb. The semantics of a novel word can be partially inferred from the linguistic context in which it occurs. The famous cases concern inferences about the meanings of verbs via syntactic bootstrapping (Gleitman, 1990). However, the effect is more general. Words that are similar in meaning tend to occur in similar linguistic contexts, providing a basis for rapidly inferring the meanings of new words (Landauer & Dumais, 1997). In short, nonword stimuli do not preclude semantic effects.

3. Semantic distance is correlated with derivational status and so both factors predict the same results for most words. For four denominals in the Kim et al. study, the two factors dissociated. Derivational status predicted that the irregular past tense would be disallowed. Semantic distance predicted that the irregular past would be preferred; for these items, the derived meaning was rated as more semantically similar to the homophonic verb than to the putative source noun. In the few cases where the two factors made different predictions, semantics trumped derivational status (see also Ramscar, 2002).[6] Thus semantic factors provide a viable unifying generalization, difficult though it may be to quantify semantic similarity via rating scales.

One final case with a similar character concerns the impact of brain injury on the use of inflectional morphology; see McClelland and Patterson (2002) for a summary. In brief, some studies of brain-injured patients have been taken as showing that the two components of the rules-and-words theory can be independently impaired by different neuropathologies, evidence that they are distinct modules with different biological substrates. McClelland and Patterson summarize research suggesting that different patterns of impairment arise from damage to subsystems supporting phonological versus semantic information. In general, semantics makes greater contributions to irregular verbs, which are more affected by neuropathologies such as semantic dementia (Patterson et al., 2006). Phonological impairments have a greater impact on regular past tenses, because regular

but not irregular past tenses are phonologically complex (Burzio, 2002), and phonologically impaired patients have difficulty producing complex words, whether they are inflected (as in *flaked*) or not (as in *strict*). Their account is convincing because it can be independently established that the patients in question have semantic versus phonological impairments, which affect many aspects of performance, not just verb morphology.

9. Summary

Despite extensive evidence that people activate multiple types of information in real time to compute contextually appropriate meanings (MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), the two-module theory holds that the structure of the language processing system is such that only certain types of information are available to processing modules at specific points in time. The subject in a past tense generation task hears the word *bake*, the recognition of which makes available information about its sound, meaning, grammatical category, and the structures in which it participates (MacDonald et al., 1994). In the two-module theory, only grammatical information enters into determining whether a word is rule governed or not. However, the system must also be structured so as to allow phonological and semantic information to be accessed once this initial decision is made. It cannot literally be true that rule application is blind to phonological properties of verbs because the rule *refers* to a phonological property of the words to which it applies—the final phoneme. Rather, the theory holds that the rule is blind to some phonological properties of words at some moments in processing. Similarly, semantics is not initially accessible, but must become available to account for various semantic influences reviewed above. This elaborate gating of different types of information over time is at odds with the highly interactive, immediate character of processing established in research on comprehension and production, including important work by Rumelhart (1977), McClelland and Rumelhart (1981), and Marslen-Wilson (1975) that predated the PDP approach (see also McClelland, Mirman, Bolger, & Khaitan, 2014).

Despite its many problems, the two-module theory is difficult to dispatch. A strong, authoritative theory that provides elegant analyses of well-chosen examples has enormous appeal, as does the intuition that language is rule governed except for exceptions that must be memorized. The theory addresses an idealized version of performance that allows outlier data to be excluded. The theory is not easy to disconfirm via critical experiments because of the loose coupling between theory and data. Whatever the value of this account, its limitations reinforce the need to consider other approaches.

10. What's the matter with modeling?

Whereas it is too difficult to disconfirm the two-module account, models like Rumelhart and McClelland's are too easy to disprove. Models are tools for investigating theoretical principles and empirical phenomena. Every model only partially instantiates the

principles in question and addresses only some of the phenomena. Every model makes simplifying assumptions about some components that enable investigations of important phenomena. Every model addresses a limited range of phenomena. People always bring more to a task than can be represented in a model of it. The fundamental axiom of computational modeling is that *every model is literally and inherently false at some level of detail.* Models are nonetheless necessary in order to determine whether proposed mechanisms can in fact give rise to phenomena of interest (McClelland, 2009; Seidenberg, 2011).

One response to this dilemma is to distinguish between theory-relevant and implementation-specific aspects of a model (MacWhinney & Leinbach, 1991). Whereas Pinker and Prince (1988) attributed the limitations of the RM model to the PDP approach rather than to that particular instantiation of it, others saw opportunities to move forward. For example, the RM model focused on continuity between regulars and irregulars, and generalization without rules. Their phonological representation was intended to capture enough about phonological structure to be able to explore these issues, in advance of a complete account of phonological acquisition and processing. The limitations of this representation and its impact on generalization in particular became apparent through analyses of subsequent models of the past tense (e.g., MacWhinney & Leinbach, 1991) and related phenomena in reading (Plaut et al., 1996; Seidenberg & McClelland, 1989). The correction of this "implementational" problem was interesting insofar as it led to a deeper understanding of properties of distributed representations that affect learning (Plaut et al., 1996). It also showed how anomalous behavior can be traced to a property of an implemented model rather than the theory it instantiated.

The same could be asked of every limitation of the RM model. Was the treatment of the U-shaped learning phenomenon fundamentally wrong or was its plausibility undercut by the way it was implemented? The question was addressed by developing models that both simulated the phenomenon more closely and provided a deeper understanding of the factors that give rise to it (Plunkett & Marchman, 1991, 1993; Rogers, Rakison, & McClelland, 2004).

As a third example, the RM model could not produce different past tenses for homophones such as *ring-ringed* (the city) versus *ring-rang* (the bell) versus *wring/wrung* (the cloth). This limitation is built-in—a model that only represents phonology must represent all three forms of *(w)ring* the same way—but is it inevitable? There are two possibilities:

1. The phenomenon is important but outside the scope of the RM model, which addressed other issues for which homophone disambiguation had no relevance. If this is correct, it should be possible to address the concern in an extended model that included the phenomenon within its scope.

2. The phenomenon is important but no extension of the RM model could ever address them because of intrinsic flaws in the approach. The only extended model that could handle these and other unaddressed phenomena would be a version of the two-module theory (Pinker & Prince, 1988).

Eventually Joanisse and Seidenberg (1999) described a model that learned about the present and past tenses of verbs in the course of performing a variety of tasks (production

and comprehension of present and past tense verbs under different conditions), using semantic and phonological information. The conjunction of semantics and phonology allows the correct computation of different past tenses for both phonologically similar but semantically different homophones like *ring (rang, ringed)*, and for phonologically different but semantically similar synonyms such as *slap*, *strike*, and *hit* (see also Cottrell & Plunkett, 1995; Woollams, Joanisse, & Patterson, 2009).[7]

These cases illustrate important characteristics of modeling relevant to the past tense and beyond. The goal of the modeling is not to merely reproduce behavior, but rather to understand how behavior arises from more basic principles. Given that goal, both the successes and failures of models are informative (McClelland, 2009; McClelland, McNaughton, & O'Reilly, 1995). The fact that a model's behavior differs from people's is not itself informative because such limitations are intrinsic to the methodology. The payoff comes from analysis of why the model behaves the way it does. This analysis may validate the model's core assumptions (e.g., the analysis of how properties of phonological representations affect performance left RM's account of generalization intact), or it may reveal fundamental limitations that lead to theoretical insights (as in McClelland et al., 1995). The same is true of a model's successes. The interest is not simply in whether the model "works," but why: whether the behavior arises from fundamental principles or is artifactual—that is, it works but for the wrong reasons.

In practice, the limitations of a model can often be successfully addressed by doing a better job with aspects of the implementation that were not initially seen as central to the phenomena in question, without changing the account of it in any significant way. Improving the phonological representation (Plaut et al., 1996) required developing a better understanding of distributed representations and learning, not a change in how phonology is computed from orthography. Adding semantics to a verb model (Joanisse & Seidenberg, 1999) involved adding unimplemented components in order to handle additional phenomena, not a change in theoretical assumptions. Such cases validate the distinction between implementational versus theory-relevant aspects of a model.

Finally, models are often implemented in ways that make learning harder than for a child, which makes direct comparisons more difficult. People always bring more to a task than can be represented in any model. Babies know far more about phonology in advance of learning the past tense than models do. There may be complementary learning systems in the brain (McClelland et al., 1995), but no past tense model has incorporated them. Models trained using algorithms such as back-propagation receive information about the correct output on every trial, but more variable, even noisy feedback, like that available to children, may lead to more robust learning (Elman & Zipser, 1988; Kalish, Rogers, Lang, & Zhu, 2011). Models are simplified in ways that maintain tractability but at the cost of excluding factors that may contribute to children's superior performance. (This issue is clearly relevant to why models often apparently require more learning trials than children.) Identifying these differences is also important. Models are not people and do not behave exactly like them. Insights arise from analyzing where and why they differ, and whether the differences are relevant to any given issue.

In summary, a variety of limitations are intrinsic to the modeling methodology, which complicate the interpretation of a model's behavior and comparisons to people. Identifying these limitations and understanding their origins and implications are an essential part of the approach (Seidenberg & Plaut, 2006).

11. All's well that ends well?

Not entirely. The RM model and accompanying critiques stimulated the development of a succession of models. There were models that addressed the limitations of earlier models; models that explored additional computational mechanisms; models that were applied to an ever-growing body of empirical phenomena. Every model was implemented somewhat differently than every other model. Every model addressed an interesting but limited range of phenomena; every model was literally and inherently false at some level. Determining whether the limitations of a model were implementational or theory relevant required more models. Past tense modeling came to resemble a game of Whack-A-Mole. Every time a model addressed one problem, another would pop up. The many revised and improved models did not lead to the development of an integrative model that addressed all of the major phenomena.[8]

Herein lies a deeper issue. Models require distinguishing between theoretically relevant versus implementation-specific properties. A model is observed to exhibit behaviors that have theoretical significance. The model inevitably deviates from human performance in some way. For proponents of these models, deviations from expected behaviors merely show that performance is affected by other, implementational factors unrelated to the core theory. We have heard this story before: It is competence-performance redux. First, (Chomsky) deviations from the idealized theory of linguistic competence were attributed to the distorting influence of performance. Then, (Pinker) deviations from the idealized theory of performance were attributed to other performance imperfections. Now, (Rumelhart and McClelland, MacWhinney and Leinbach) deviations from a computational model are attributed to implementational factors. We all take credit for successes and outsource the failures.

Given these similarities, why bother modeling? It's difficult, time-consuming, and the models never "work." If the goal of the modeling is to arrive at broader generalizations that do not depend on the details of any single implemented model, and it is the principles that the models instantiate that are theoretically relevant, and it is too difficult to distinguish the theoretical wheat from the implementational chaff, let's skip the models entirely. Instead, we could aim to clearly articulate a principled theory of how something like the past tense works, using many different types of evidence and arguments to increase our confidence that the theory is correct. Having established a basis for believing that the theory is likely to be correct, we could simply assign appropriate behaviors to it such that, for example, application of the past tense rule is sensitive to phonology where it should be and not where it should not. Filling in the mechanistic details is unimportant. Given the plausibility of the high-level theory—and having taken any competing theory

off the table—we can safely assume that this is the system that is actually implemented, somehow.

We have just described Pinker's approach to the past tense. There is a high-level theory, mainly motivated by linguistic considerations (e.g., the putative need to distinguish grammar from lexicon or the intrinsically hierarchical structure of lexical representations). These core theoretical claims are supported by telling examples and specific findings. There is little need to develop a more mechanistic account because these analyses establish both the credibility of the theory and the inadequacy of competing accounts (e.g., PDP), which lack these properties, and seem problematic in other respects.

We need a moment to imagine what a PDP version of this approach might be like. Core theoretical principles are stated (e.g., Rumelhart, McClelland & the PDP Research Group, 1986), mainly motivated by computational considerations and general properties of brain and behavior. We describe telling examples of phenomena that seem to support the principles (e.g., graceful degradation, computation of best fits, interactivity). We describe informal models of phenomena such as the past tense, perhaps illustrating them as in the box-and-arrow modeling era (Patterson, Marshall, & Coltheart, 1985). We ascribe plausible behaviors to a model, such as the generation of irregularizations like *glang*. This model works perfectly! Filling in the details of how specific behaviors arise at a mechanistic level is unnecessary and we avoid the fog of implementation entirely.

This will not do. As a theory makes closer contact with a broader range of data, further elaboration is necessary, even at this informal level. Pinker's 1999 book captures this reality well. The story opens with strong, clear assertions about core properties of the system (language, the past tense). We are drawn in by the theory's elegant account of illustrative phenomena. There follows a succession of modifications, emendations, and qualifications. These developments, some of which are highly technical, yield a far more complex theory. It becomes increasingly difficult to track if all the parts fit together. It would take a computational model to tell, but there isn't one.

Computational modeling performs the several essential functions we described earlier (see also Seidenberg & Plaut, 2006). Perhaps the most basic function is to provide a strong test of the adequacy of a theoretical proposal: whether proposed mechanisms can in fact give rise to target phenomena; whether a mechanism introduced to address one phenomenon has an impact on explanations of other phenomena; whether the account of how knowledge is acquired is consistent with the account of how it is used; whether the theory makes different empirical predictions than any other approach. The statement of general principles in the PDP volumes was coupled with applications to diverse phenomena (including the past tense). To abandon explicit modeling in favor of a return to box-and-arrow diagrams would be a step backward (see Seidenberg, 1988, to reconnect with this earlier era).

Finally, the analogy we made earlier, equating Chomsky's competence theory, Pinker's competence theory of performance, and the reflexive ascription of a model's failures to the implementation, is specious. A model's behavior is determined by the implementation. Identifying the causes of behavior, good or bad, is a decidable issue. It

isn't merely assumed that all unwanted behaviors are due to unimportant implementational details; the causes of the behaviors have to be identified. Determining why a model behaves as it does is part of the approach. One can experiment with an implemented model to determine how elements of the implementation contribute to the results. Formal or quantitative analyses of networks can provide greater insight about a model's behavior. Comparing models of related phenomena that differ in implementational detail serves a similar function. Such analyses clarify a model's contributions and establish directions for additional research with the goal of identifying generalizations that transcend the details of any single model. These options are not available if the "model" is unimplemented.

In summary, computational models play an important role in developing explanatory theories of behavior and its brain bases. This observation applies to models of the past tense, to the class of PDP models, and to other types of quantitative and computational modeling. The alternative style of theorizing is often highly technical (e.g., in the description of linguistic structures) but lacks details about the mechanisms that underlie learning and behavior. The validity of the theory rests on establishing its plausibility rather than developing core assumptions fully enough to be testable. For example, the two-module theory holds that children learn linguistic rules, but it lacks an account of how a rule such as the past tense could be acquired, under the conditions that obtain in the course of learning a language. The failure to address the mechanisms that underlie behavior creates enormous gaps in understanding.

12. Looking past the past

Some 25 years later, there is still little agreement about the outcome of the past tense debate. For nativists, the Pinker and Prince critique of Rumelhart and McClelland was a replay of Chomsky's (1959) critique of Skinner. From this perspective, both episodes involved turning back attempts to reduce language to "associations."

On the PDP side, the past tense debate was a replay of a different event, Minsky and Papert's (1969) critique of perceptrons. From this perspective, both episodes involved authoritative but ultimately mistaken overstatements about the limitations of neural networks based on analyses of specific cases. Perceptrons did not exhaust the range of neural network architectures; the Rumelhart and McClelland model did not exhaust the range of PDP models of the past tense. Both critiques prematurely shunted attention and resources away from further research.

Perhaps the main legacy of Rumelhart and McClelland's work is the formulation and continuing development of a theory that situates the past tense in the broader context of a lexical system that acquires and represents several types of information (mainly semantics, phonology, orthography) and the relations between them. This lexical network is used in comprehending (listening, reading) and producing (speaking, writing) language (Joanisse & Seidenberg, 1999). The past tense is part of this system, rather than part of "grammar." This lexical system develops in learning to use language to perform naturalistic

tasks such as reading and listening, but, like people, it also supports performing the laboratory tasks—like the WUG task—used to systematically investigate its properties. The covering generalizations are about lexical processing, not the past tense or inflectional morphology.

Much research within this framework has addressed word reading. Reading is important to study because of its cultural significance and cognitive complexity, but it is thought to have little linguistic interest because writing was invented well after speech emerged in humans. It is therefore notable that reading shares so many characteristics with the past tense, which was said to exemplify core properties of language, a biological capacity. The reading models (Harm & Seidenberg, 1999; Plaut et al., 1996; Seidenberg & McClelland, 1989) initially focused on the quasiregular correspondences between spelling and sound in English. As with the past tense, a network that encoded the consistency of the mappings between codes could generate both "rule-governed" cases and "exceptions," while capturing the overlap between the two and supporting generalization. Later reading models (Harm & Seidenberg, 2004; Woollams et al., 2007) investigated issues that closely paralleled research on verb morphology: the contributions of semantic and phonological information in generating well-formed output; "division of labor" between components of the system in supporting skilled performance; acquisition of these components and the bases of developmental impairments; differential impact of damage to semantic and phonological components of the lexical system; and other topics. There are close parallels because both inflectional morphology and word reading are expressions of the same lexical processing system. This account correctly predicts that damage to a major component of this system, such as semantics, should affect both inflectional morphology and word reading, as well as other tasks that depend on this code (Patterson et al., 2006).

A major feature of these systems is their capacity to develop internal representations of structural regularities that are discovered in the course of learning. The models of reading aloud learned to perform the task efficiently by discovering correlations between orthography and phonology, which in English occur over units of various sizes, ranging from individual letters to letter combinations to entire words (Treiman, Kessler, & Bick, 2003; Zevin & Seidenberg, 2006). The same process occurs in writing systems employing other types of units, such as Chinese (Yang et al., 2009). The discovery of such "islands of reliability" (Albright & Hayes, 2003) occurs in the course of acquiring a vocabulary of thousands of words, as described by Seidenberg and McClelland (1989). The components of a word such as *bag* that are relevant to pronunciation are uncovered as the network learns overlapping words such as *bat*, *bad*, and *beg*. The same principles underlie the emergence of morphemes, which are "islands of reliability" in the mappings between form (orthography, phonology) and meaning (Seidenberg & Gonnerman, 2000). Morphology is quasiregular because the degree to which these codes converge varies. Classical morphemes (e.g., the *think* in *think*, *thinker*, and *rethink*) make relatively consistent contributions across words, but often the convergences are only partial. Although *write-rewrite* and *turn-return* are structurally similar, the former have greater semantic and phonological overlap than the latter. *Bakery* is related to *bake* and *cannery* to *can* but

there is no *groce* in *grocery*. All the berry words (*blueberry, blackberry, strawberry, cranberry* et al.) are structurally similar, but whereas blueberries are blue and blackberries are black(ish), *straw* is semantically unrelated to *strawberry* and *cran* is not a word (see Seidenberg & Gonnerman, 2000, for other examples and discussion, and Plaut & Gonnerman, 2000, for models of how these processes play out in typologically different languages such as Hebrew and English). Language users develop sensitivity to the internal structure of words, but the units are graded rather than discrete and their interpretations are context-dependent (Gonnerman, Seidenberg & Anderson, 2007).

The past tense and other aspects of inflectional morphology are also governed by these mechanisms. The internal structure of a verb like *walked* is discovered through exposure to other words containing *walk* (e.g., *walks*, *walking*) or *–ed* (other regular past tense verbs). The language user learns that *walked* consists of two subunits that make consistent contributions to different words, whereas *took* does not (see Willits, Seidenberg & Saffran, 2014, for evidence concerning the special role of *–ing* in this process). *Slept* is an intermediate case because it is pulled in two directions: it has the standard past tense inflection but *slep* is degenerate. The same forces cause /prest/ *(pressed)* to develop strong internal structure but less so for the phonologically similar /tʃest/ *(chest)*.

These extensions of the RM approach have generated their own controversies. The view that morphemes reflect the covergence of form and meaning contrasts with theories in which morphology is an abstract level of linguistic representation, with discrete units defined structurally, independent of semantics (Marlsen-Wilson & Tyler, 2007). The primary psycholinguistic evidence for this view comes from studies of morphological priming (e.g., Rastle et al., 2000; Rastle, Davis & New, 2004) that manipulate the degree to which prime and target pairs overlap in form and meaning. The critical finding is that priming is observed for pairs such as *blaze-blazer* and *seed-seedy* that are related in form but not meaning, compared to relevant controls. Thus, whereas the Seidenberg and Gonnerman (2000) theory holds that morphology *emerges* from correlations among codes, the alternative theory treats correlations between form and meaning as a *confound* that must be controlled through stimulus selection in order to reveal the true nature of morphological representation. In fact the strong correlations between form and meaning in English make it difficult to manipulate these properties independently and existing evidence is contradictory.[9]

Beyond these direct extensions of their work, Rumelhart and McClelland's ideas continue to exert broad influence via their assimilation into the theoretical mainstream, particularly in the study of language. The view that language should be understood as a quasiregular system with graded statistical structure is now implicit in studies of most aspects of language use. In language acquisition, for example, there has been a turn toward examining the role of "statistical learning," facilitated by the development of methods suitable for obtaining relevant data from young subjects (e.g., Saffran, Aslin, & Newport, 1996). Structures that support language use develop through statistical learning over the input to the child, constrained by the state of the child's developing perceptual, motor, and memory systems (Seidenberg, 1997). There is a robust empirical enterprise devoted to identifying the properties of statistical learning, such as the kinds of units it operates over (Gerken, Wilson, & Lewis, 2005; Willits, Seidenberg, & Saffran, 2014),

whether learning occurs over nonadjacent elements (Newport & Aslin, 2004), whether it occurs over different types of structures simultaneously (Sahni, Seidenberg, & Saffran, 2010) and so on. In a parallel development, studies of adult performance now focus on the use of language statistics in online processing (MacDonald & Seidenberg, 2006), facilitated by the availability of large language corpora. This research is making important progress toward understanding how we manage to acquire and use language. It is clear, as Rumelhart and McClelland's work suggested, that far more information can be derived from the rich yet noisy and variable input to the child than had been concluded from arguments about the "poverty of the stimulus."

Although studies of "statistical learning" have energized language acquisition research, the methods are more in keeping with traditional experimental psychology than guided by computational models. Little of this work is explicitly linked to Rumelhart and McClelland (1986), to the PDP approach, or to computational models of learning. This disconnection is especially striking because so much of the research concerns statistical regularities over sequences of elements such as syllables and words, the learning of which was the focus of Elman's (1990) simple recurrent network (SRN) approach. The verb and reading models focused on people's knowledge of words and simple relations between words (e.g., present and past tense). However, words occur in sentences that exhibit additional statistical regularities that reflect how they combine to convey utterance meaning. Elman's model suggested that such regularities, and thus much of the structure of language, could be encoded by a sequential processing mechanism that learned via predicting upcoming input.

Many studies of language acquisition address issues closely related to the Elman work, such as the use of sequential statistics to identify the boundaries between words in continuous speech (Saffran etal., 1996) and the induction of grammatical categories and meanings of words from distributional regularities (Reeder, Newport & Aslin, 2013). Similarly, Elman's approach presaged current work in sentence comprehension, in which prediction is seen as a key mechanism (Levy, 2008). Here too the empirical literature reflects the adoption of core PDP ideas, with minimal linkage to the computational models themselves.

The adoption of PDP concepts while disavowing the computational models is quite striking as reflected in this observation by Newport (2010), who describes "the movement in many parts of psycholinguistics from rules to connectionism to statistical learning" (p. 369). What is missing from this post-connectionist research is a theory of what statistical learning *is.* Studies show that infants are sensitive to particular statistical regularities. Languages exhibit innumerable such regularities, most of which turn out to be trivial. Addressing why some regularities are picked up and not others, why statistics are computed over particular units, and how such units emerge and change over development requires an explicit, developmental theory of statistical learning which, taken with other facts about children's capacities and experiences, could explain the range of outcomes that are observed. Cognitive Bayesian theories do not serve these functions for reasons discussed by McClelland et al. (2010).

The Rumelhart and McClelland model and the work it inspired had such a theory as a goal. The research was, above all else, about the mechanisms of learning and development

(see also Elman et al., 1996; Munakata & McClelland, 2003). The models instantiated specific hypotheses about the data reduction and regularization that occur in learning from noisy, variable, quasiregular input such as language. This emphasis on an explicit, mechanistic account of learning and developmental change has been lost in the transition from "connectionism" to "statistical learning." The result is a collection of findings in search of a general theory.

Similarly, Rumelhart and McClelland's emphasis on learning and development suggested that the core issues of language acquisition, skilled performance, and brain bases could be unified within a common computational framework. "Acquisition" and "skilled performance" reflect the performance of a network at different points in development. The claim is important but difficult to assess without implementing explicit models that transition from learning language to using it. Such models do not appear to be on the horizon in studies of language acquisition or skilled performance.

In the end, there is irony in the fact that, whereas Rumelhart and McClelland's main ideas have survived and pervade current thinking about language well beyond the past tense, the use of the methods that were essential to their development is more limited. The detailed critiques of specific models, overly focused on unimportant properties, managed to undermine confidence in the methodology in some circles. Rumelhart and McClelland's ideas, deeper than any single model, have proved more difficult to set aside.

Acknowledgments

The authors are grateful to Karalyn Patterson for her insights about the past tense, for comments on an earlier version of the manuscript, and for retrieving some information the authors had suppressed. Maryellen MacDonald greatly clarified several linguistic phenomena; both she and Jay McClelland provided especially helpful comments on earlier versions. We also thank the editors of the issue for their patience, and the reviewers for their comments and suggestions.

Notes

1. Our review focuses on the characterization of the past tense and related inflectional phenomena, and behavioral evidence concerning acquisition and skilled performance. The neurobiological evidence from neuroimaging and studies of brain-injured patients is discussed elsewhere (e.g., McClelland and Patterson, 2002; Woollams et al., 2007; Patterson et al., 2006; Marslen-Wilson and Tyler, 2007; Desai et al., 2006).
2. Pioneers such as Lenneberg (1967) attempted to study the brain bases of language more directly, an approach that eventually flourished with the development of modern methods such as neuroimaging, coupled to computational modeling.

3. The late John McCarthy (Stanford AI) wrote, "In 1965, the Russian mathematician Alexander Kronrod said, 'Chess is the *Drosophila* of artificial intelligence.' However, computer chess has developed much as genetics might have if the geneticists had concentrated their efforts starting in 1910 on breeding racing *Drosophila*. We would have some science, but mainly we would have very fast fruit flies." McCarthy (1990).
4. The major cross-linguistic evidence supporting the two-module theory was supplied by studies of the German plural. Rules are said to be the basis for generalization: Given a novel form, people default to applying the rule. Marcus et al. (1995) argued that German has rule-governed plurals but unlike English, they are less frequent than the several types of irregular plurals in that language. Since PDP models are only sensitive to pattern frequencies, they argued, the models cannot produce the lower frequency regular forms as generalizations. The accuracy of their characterization of plural formation in German was challenged, however (Hahn & Nakisa, 2000; Penke & Krause, 2002). Plural formation in German depends on a conjunction of semantic and phonological factors. As in English, generalization does not rigidly conform to the rule, producing irregularized plurals, analogous to English nonce verbs such as *splung*.
5. The word refers to a single object, but there is some plural semantics in so-called bifurcate pluralia tanta such as *scissors, pants,* and *trousers*.
6. According to Kim et al. (1991), data that deviate from their predictions could also be due to subjects' "uncertainty" about the derivational status of a word. Such uncertainty is widespread: it is shared by experts for whom the derivational status of many of Kim et al.'s stimuli is also unclear (Ramscar, 2002). Thus neither the information to be acquired nor the means by which children could acquire it are known. These considerations accord with other evidence that derivational status is not a salient factor in determining the past tenses of novel verbs.
7. Pinker and Ullman (2002) dismissed this model as an implementation of their two-mechanism theory because it used a simplified representation in which the semantics of each word was represented by a single unit, which they equated with the lexical entries in their theory. However, the same results obtain using distributed semantic representations (Woollams et al., 2009). A further challenge was presented by the identification of brain-injured patients with unusual deficit patterns. Such case studies provide important information and often drive research in interesting new directions, but they are also open to multiple interpretations (see references in note 1). As an example, Miozzo (2003) described a patient who could generate past tenses for regular verbs but was impaired on irregulars. This pattern can arise from semantic damage in models such as Joanisse and Seidenberg (1999). However, this patient's semantic knowledge was relatively intact, as indicated by performance on a test of ability to match spoken words to pictures. Hence the patient's performance seems to decisively refute the model. The key additional fact is that the patient was severely anomic—poor at naming pictured objects—indicating an impairment in the use of semantics to generate phonological output. Accord-

ing to Joanisse and Seidenberg (1999), generating irregular past tenses requires greater reliance on exactly this process—the semantics→phonology side of the model—than does generating regular past tenses (see Figure 5, McClelland & Patterson, 2002). Thus, the patient's performance follows from properties of the model rather than contradicting it.

8. Pinker's (1999) characterization of this history is staggeringly different: "After Alan Prince and I took apart the pattern associator [RM] model, the linguists breathed a sigh of relief because they thought they didn't have to learn neural network modeling after all, and the connectionists dropped it like a hot potato." (p. 117).
9. Research in this area is ongoing. The masked priming studies do not consider phonological properties of words, which vary across items such as PAINT-PAINTER, REBEL-REBELLIOUS, and COMPETE-COMPETENCE. Tasks differ in their sensitivity to semantic information. Experiments using these types of stimuli with other tasks have yielded different results (e.g., Duñabeitia, Kinoshita, Carreiras, & Norris, 2011; Frisson, Bélanger & Rayner, 2014).

References

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, *90*, 119–161.

Albright, A., & Hayes, B. (2006). Modeling productivity with the gradual learning algorithm: The problem of accidentally exceptionless generalizations. In G. Fanselow, C. Fery, M. Schlesewsky, & R. Vogel (Eds.), *Gradience in grammar: Generative perspectives* (pp. 185–204). Oxford, England: Oxford University Press.

Alegre, M. A., & Gordon, P. (1996). Red rats exposes recursion in children's word formation. *Cognition*, *60*, 65–82.

Allen, J., & Seidenberg, M. S. (1999). Grammaticality judgment and aphasia: A connectionist account. In B. MacWhinney (Ed.), *The emergence of Language*. Hillsdale, NJ: Erlbaum.

Baker, C. L. (1981). *The logical problem of language acquisition*. Cambridge, MA: MIT Press.

Berent, I., & Pinker, S. (2007). The dislike of regular plurals in compounds: Phonological familiarity or morphological constraint? *The Mental Lexicon*, *2*, 129–181.

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, 150–177.

Bickerton, D. (1984). The language bioprogram hypothesis. *The Behavioral and Brain Sciences*, *7*, 173–222.

Bird, H., Lambon Ralph, M. A., Seidenberg, M. S., McClelland, J. L., & Patterson, K. (2003). Deficits in phonology and past-tense morphology: What's the connection? *Journal of Memory and Language*, *48*, 502–526.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart, & Winston.

Burzio, L. (2002). Missing players: Phonology and the past-tense debate. *Lingua*, *112*, 157–199.

Bybee, J., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, *22*, 381–410.

Bybee, J., & Moder, C. L. (1983). Morphological classes as natural categories. *Language*, *59*, 251–270.

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, *58*, 265–289.

Chien, Y., & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*, *1*, 225–295.

Chomsky, N. (1959). A review of B. F. Skinner's *Verbal behavior*. In N. Block (Ed.), *Readings in philosophy of psychology*, Volume 2. Cambridge, MA: Harvard University Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, the Netherlands: Foris.

Chomsky, N. (1985). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Crain, S. (1991). Language acquisition in the absence of experience. *The Behavioral and Brain Sciences*, *14*, 597–650.

Daugherty, K. G., MacDonald, M. C., Petersen, A. S., & Seidenberg, M. S. (1993). Why no mere mortal has ever flown out to center field but people often say they do. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Daugherty, K. G., & Seidenberg, M. S. (1994). Beyond rules and exceptions: A connectionist approach to English morphology. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules*. Amsterdam: John Benjamins.

Desai, R., Conant, L. L., Waldron, E., & Binder, J. R. (2006). fMRI of past tense processing: The effects of phonological complexity and task difficulty. *Journal of Cognitive Neuroscience*, *18*, 278–297.

Duñabeitia, J. A., Kinoshita, S., Carreiras, M., & Norris, D. (2011). Is morpho-orthographic decomposition purely orthographic? Evidence from masked priming in the same–different task. *Language and Cognitive Processes*, *26*, 509–529.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Elman, J. L., & Zipser, D. (1988). Discovering the hidden structure of speech. *Journal of the Acoustical Society of America*, *83*, 1615–1626.

Fodor, J., & Pylyshyn, Z. (1988). Connectionsim and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.

Frisson, S., Bélanger, N. N., & Rayner, K. (2014). Phonological and orthographic overlap effects in fast and masked priming. *The Quarterly Journal of Experimental Psychology*, published on-line, July 31, 2014

Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*, 249–268.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3–55.

Gold, E. M. (1967). Language identification in the limit. *Information and Control*, *10*, 447–474.

Gordon, P. (1985). Level-ordering in lexical development. *Cognition*, *21*, 73–93.

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.

Hahn, U., & Nakisa, R. C. (2000). German inflection: Single route or dual route? *Cognitive Psychology*, *41*, 313–360.

Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry*, *4*, 3–16.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720.

Harris, C. L. (1992). Understanding English past-tense formation: The shared meaning hypothesis. In *Proceedings of the 14th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Haskell, T. R., MacDonald, M. C., & Seidenberg, M. S. (2003). Language learning and innateness: Some implications of compounds research. *Cognitive Psychology*, *47*, 119–163.

Huang, Y. T., & Pinker, S. (2010). Lexical semantics and irregular inflection. *Language and Cognitive Processes*, *25*, 151.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Science USA*, *96*, 7592–7597.

Kalish, C. W., Rogers, T. T., Lang, J., & Zhu, X. (2011). Can semi-supervised learning explain incorrect beliefs about categories? *Cognition*, *120*, 106–118.

Kim, J. J., Pinker, S., Prince, A., & Prasada, S. (1991). Why no mere mortal has ever flown out to center field. *Cognitive Science*, *15*, 173–218.

Kiparsky, P. (1982). *Explanation in phonology*. Dordrecht, the Netherlands: Foris.

Lachter, J., & Bever, T. G. (1988). The relation between linguistic structure and theories of language learning: A constructive critique of some connectionist learning models. *Cognition*, *28*, 195–247.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177.

Lightfoot, D. (1982). *The language lottery: Toward a biology of grammars*. Cambridge, MA: MIT Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction processes in lexical and sentence processing. In M. Traxler, & M. Gernsbacher (Eds.), *Handbook of psycholinguistics research* (2nd ed.) (pp. 581–612). New York: Academic Press.

MacWhinney, B. (Ed.) (1999). *The emergence of language*. Mahwah, NJ: Erlbaum.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, *40*, 121–153.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, *29*, 189–256.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, J. T., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*, 1–165.

Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research*, *2*, 175–199.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, *189*, 226–228.

Marslen-Wilson, W. D., & Tyler, L. K. (2007). Morphology, language and the brain: The decompositional substrate for language comprehension. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*, 823–836.

McCarthy, J. (1990). Chess as the Drosophila of AI. In T.A. Marsland & J. Schaeffer (Eds.), *Computers, chess, and cognition* (pp. 227–237). New York: Springer.

McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science*, *1*, 11–38.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to understanding cognition. *Trends in Cognitive Sciences*, *14*, 348–356.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, *6*, 465–472.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.

McClelland, J. L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, *4*, 287–335.

McCloskey, M. (1991). Networks and theories: The place of connectionism in cognitive science. *Psychological Science*, *2*, 387–395.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation*. Vol. *24* (pp. 109–165). New York: Academic Press.

Minsky, M., & Papert, S. (1969). *Perceptrons: An introduction to computational geometry*. Cambridge, MA: MIT Press.

Mirković, J., MacDonald, M. C., and Seidenberg, M. S. (2005). Where does gender come from? Evidence from a complex inflectional system. *Language and Cognitive Processes*, *20*, 139–167.

Miozzo, Michele (2003). On the processing of regular and irregular forms of verbs and nouns: Evidence from neuropsychology. *Cognition*, *87*, 101–127.

Mirković, J., Seidenberg, M. S., & Joanisse, M. F. (2011). Rules versus statistics: Insights from a highly inflected language. *Cognitive Science*, *35*, 638–681.

Munakata, Y., & McClelland, J. L. (2003). Connectionist models of development. *Developmental Science*, *6*, 413–429.

Nakisa, R. C., Plunkett, K., & Hahn, U. (2001). A cross-linguistic comparison of single and dual-route models of inflectional morphology. In P. Broeder & J. Murre (Eds.), *Models of language acquisition: Inductive and deductive approaches*. Cambridge, MA: MIT Press.

Newport, E. L. (2010). Plus or minus 30 years in the language sciences. *Topics in Cognitive Science*, *2*, 367–373.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162.

Patterson, K., Lambon Ralph, M. A., Jefferies, E., Woollams, A., Jones, R., Hodges, J. R., & Rogers, T. T. (2006). "Presemantic" cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience*, *18*, 169–183.

Patterson, K., Marshall, J. C., & Coltheart, M. (Eds.) (1985). *Surface dyslexia*. Hillsdale, NJ: Erlbaum.

Penke, M., & Krause, M. (2002). German noun plurals: A challenge to the dual-mechanism model. *Brain and Language*, *81*, 303–311.

Perfors, A., Tenenbaum, J. B., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, *118*, 306–338.

Piattelli-Palmarini, M. (1994). Ever since language and learning: Afterthoughts on the Piaget-Chomsky debate. *Cognition*, *50*, 315–346.

Pinker, S. (1991). Rules of language. *Science*, *253*, 530–535.

Pinker, S. (1994). *The language instinct*. New York: Morrow.

Pinker, S. (1997). Words and rules in the human brain. *Nature*, *387*, 547–548.

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York: Basic Books.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*, 73–193.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*, 456–463.

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, *15*, 445–485.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

Plunkett, K., & Marchman, V. A. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, *38*, 43–102.

Plunkett, K., & Marchman, V. A. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, *48*, 21–69.

Prasada, S., & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, *8*, 1–56.

Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. London: Blackwell.

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology*, *45*, 45–94.

Ramscar, M., & Dye, M. (2011). Learning language from the input: Why innate constraints can't explain noun compounding. *Cognitive Psychology*, *62*, 1–40.

Ramscar, M., Dye, M., & Hübner, M. (2013). When the fly flied and when the fly flew: How semantics affect the processing of inflected verbs. *Language and Cognitive Processes*, *28*, 468–497.

Rastle, K., Davis, H., & New, B. (2004). The broth in my brother's brothel: Morphoorthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*, 1090–1098.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, *15*, 507–537.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, *97*, 285–308.

Reeder, P. A., Newport, E. L., & Aslin, R. N. (2013). From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cognitive psychology*, *66*, 30–54.

Rogers, T. T., Rakison, D., & McClelland, J. L. (2004). U-shaped curves in development: A PDP approach. *Journal of Cognition and Development*, *5*, 137–145.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI* (pp. 573–603). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models* (pp. 216–271). Cambridge, MA: MIT Press.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (Eds.) (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations*. Cambridge, MA: MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Sahni, S. D., Seidenberg, M. S., & Saffran, J. R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development*, *81*, 727–736.

Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Seidenberg, M. S. (1988). Cognitive neuropsychology and language: The state of the art. *Cognitive Neuropsychology*, *5*, 403–426.

Seidenberg, M. S. (1993). Connectionist models and cognitive theory. *Psychological Science*, *4*, 228–235.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, *275*, 1599–1604.

Seidenberg, M. S., & Gonnerman, L. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, *4*, 353–361.

Seidenberg, M. S. (2011). Reading in different writing systems: One architecture, multiple solutions. In P. McCardle, J. Ren, & O. Tzeng (Eds.), *Dyslexia across languages: Orthography and the gene-brain-behavior link*. Baltimore, MD: Paul H. Brookes Publishing.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, *23*, 569–588.

Seidenberg, M. S., MacDonald, M. C., & Haskell, T. R. (2007). Semantics and phonology constrain compound formation. *The Mental Lexicon*, *2*, 287–312.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Seidenberg, M. S., & Plaut, D. C. (2006). Progress in understanding word reading: Data fitting versus theory building. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 25–49). Hove, UK: Psychology Press.

Siegel, D. (1979). *Topics in English morphology*. New York: Garland.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*, 1–29.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Treiman, R., Kessler, B., & Bick, S. (2003). Influence of consonantal context on the pronunciation of vowels: A comparison of human readers and computational models. *Cognition*, *88*, 49–78.

Ullman, M. T., Corkin, S., Coppola, M., Hicock, G., Growdon, J. H., Koroshetz, W. J., & Pinker, S. (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. *Journal of Cognitive Neuroscience*, *9*, 266–276.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, *106*, 23–79.

Willits, J., Seidenberg, M. S., & Saffran, J. (2014). Distributional structure in language: Contributions to difficulty differences in infant word recogntion. *Cognition*, *132*, 429–436.

Woollams, A. M., Joanisse, M., & Patterson, K. (2009). Past-tense generation from form versus meaning: Behavioural data and simulation evidence. *Journal of Memory and Language*, *61*, 55–76.

Woollams, A., Lambon Ralph, M. A., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review*, *114*, 316–339.

Zevin, J. D., & Seidenberg, M. S. (2006). Consistency effects and individual differences in nonword naming: A comparison of computational models. *Journal of Memory and Language*, *54*, 145–160.