Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 7592–7597, June 1999
Psychology

Impairments in verb morphology after brain injury: A connectionist model

MARC F. JOANISSE AND MARK S. SEIDENBERG*

Neuroscience Program, University of Southern California, Los Angeles, CA 90089-2520

Communicated by Richard F. Thompson, University of Southern California, Los Angeles, CA, April 9, 1999 (received for review September 30, 1998)

ABSTRACT The formation of the past tense of verbs in English has been the focus of the debate concerning connectionist vs. symbolic accounts of language. Brain-injured patients differ with respect to whether they are more impaired in generating irregular past tenses (TAKE–TOOK) or past tenses for nonce verbs (WUG–WUGGED). Such dissociations have been taken as evidence for distinct "rule" and "associative" memory systems in morphology and against the connectionist approach in which a single system is used for all forms. We describe a simulation model in which these impairments arise from damage to phonological or semantic information, which have different effects on generalization and irregular forms, respectively. The results provide an account of the bases of impairments in verb morphology and show that these impairments can be explained within connectionist models that do not use rules or a separate mechanism for exceptions.

Understanding the nature of language and its biological basis is a central issue in cognitive neuroscience. In the classical generative approach (1, 2), language is characterized in terms of a domain-specific form of knowledge representation called grammar. Grammar specifies how language is structured at multiple levels of representation using rules, constraints, and other formal symbolic devices. Generative grammar has been a prominent part of the symbolic approach to the study of cognition for many years, providing a framework for investigating important issues about the structure, acquisition, and use of language (3).

An alternative view of language has emerged from the connectionist approach, in which human capacities are understood in terms of the properties of artificial neural networks (4–6). Behavior in such systems, as in the brain, arises from interactions among large numbers of simple processing units. Acquiring a skill or type of knowledge involves adjusting the weights on connections between units, which govern patterns of activity in the network, on the basis of experience, rather than learning rules or setting language-specific parameters (7). In this view, language involves the same principles of knowledge representation, learning, and processing as other aspects of cognition. Grammars are idealized characterizations of some aspects of the behavior of the network structure actually responsible for behavior (4). This approach has attracted broad interest in the cognitive and neurosciences and has begun to provide accounts of the acquisition, representation, and use of knowledge in many domains and their neural bases.

The debate over these contrasting accounts of language has centered on a seemingly minor aspect of language, the past tense of verbs in English. The past tense has been taken to be a paradigmatic linguistic subsystem exhibiting fundamental properties of language (4, 8, 9). Grammatical features of verbs such as tense, aspect, and subject agreement are marked inflectionally (e.g., BAKE–BAKES–BAKED–BAKING). In the classical theory, inflections are generated by rules that are part of the morphological component of grammar. The past-tense rule adds the morpheme that is spelled –ED, the pronunciation of which is conditioned by the final phoneme in the base verb: if the final phoneme is a voiceless consonant, a /t/ is added (e.g., BAKED); if it is a voiced consonant or a vowel, /d/ is added (e.g., BARED); if it is an alveolar stop (/t/ or /d/), an unstressed vowel as well as /d/ are added (e.g., BAITED). The classic evidence for such rules is provided by the ability to generalize: when a new verb, such as FAX, comes into the language, speakers agree that the past tense is FAXED. Children acquire this ability to generalize at a young age (10).

There are about 180 exceptions to the past tense rule, such as SLEEP–SLEPT and TAKE–TOOK. Most aspects of language have this quasiregular character: they appear to be rule-governed but admit exceptions that deviate from the rules differing degrees (11). In traditional linguistic theories, the idea that language is rule-governed is maintained by treating irregulars separately; for example, they are thought to be learned by rote and "listed" in memory (12). This yields a dual-mechanism theory in which the rules and exceptions are governed by different principles. That these are distinct subsystems or modules is taken as an important discovery about language (8, 9).

Our approach to these phenomena focuses on people's knowledge of words and the tasks they perform using this knowledge. The representations of words include specifications of phonology (sound), semantics (meaning), and orthography (spelling), which are used in performing the primary tasks of comprehension (mapping from phonology to semantics) and production (mapping from semantics to phonology). However, this knowledge also is used to perform other tasks, such as pronouncing written words aloud and generating related forms, such as past tenses. The theoretical problem is to understand how people acquire these capacities; connectionist models provide a useful tool in this regard.

In networks using distributed representations, the same representational, learning, and processing principles govern all items. Thus, the approach does not embody the categorical distinction between rule-governed (regular) and exception (irregular) that is central to the traditional theory. Using a single mechanism to generate all forms is intended to capture the fact that the two types of forms are not categorically different; they share structure in two respects. First, there are systematic correspondences between the present and past tense forms of both regular and irregular verbs. For example, in both BAKE–BAKED and TAKE–TOOK, the past tense retains the onset and coda of the present tense (13). The degree of overlap varies across verbs, but there is some overlap for almost all past tense alternations. Second, there are similarities between regular and irregular past tense forms. For example,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: AD, Alzheimer's disease; AL, anterior-lesion aphasics; PA, posterior-lesion aphasics; PD, Parkinson's disease.

*To whom reprint requests should be addressed. e-mail: marks@gizmo.usc.edu.

CREPT and SLEPT, which are nominally irregular, end in /t/, which is the regular past tense inflection in CROPPED, STEPPED, and many other regular forms (14). In addition, there are subregularities among pools of irregular verbs, such as SING–SANG/RING–RANG and BLOW–BLEW/THROW–THREW/GROW–GREW (15). The network approach shows how these partial regularities can be picked up in the course of learning and encoded by the weights along with the regularities characteristic of the past tense rule.

Subsequent to Rumelhart and McClelland's (4) pioneering work, Plunkett & Marchman (16), Daugherty & Seidenberg (17), MacWhinney and Leinbach (18), and many others described models that learned the past tenses of both regular and irregular forms, using a common set of weighted connections between units. Such models differ from the standard approach insofar as they do not use explicit representations of rules or a separate mechanism for the exceptions.

The past tense debate has continued for about 10 years without resolution because most of the behavioral data can be accommodated by both theories (19). Attention has recently turned to evidence concerning underlying brain mechanisms (19–21). Ullman *et al.* (22) reported data concerning impairments in past tense formation in brain-injured patients. Patients with Parkinson's disease (PD) and a patient with an anterior lesion (AL) had greater difficulty generating past tenses for nonwords compared to irregulars. In contrast, patients with Alzheimer's disease (AD) or posterior aphasia (PA) exhibited the opposite pattern, with more difficulty on irregular than nonwords. This type of "double dissociation" is standardly interpreted as evidence for separate mechanisms that can be independently impaired by neuropathology (23).

Ullman *et al.* (22) proposed that the rule mechanism is a type of procedural knowledge located in the left inferior frontal cortex (including Broca's area) that is impaired by lesions to this region (AL; ref. 24) or to areas of the basal ganglia which project onto it (PD; ref. 25). The irregular mechanism is a type of declarative knowledge represented in Wernicke's area, which can be impaired by either brain lesion (PA; ref. 24) or neurodegeneration (AD; refs. 26 and 27). Hence, the study provided evidence for the brain bases of the two proposed modules and seemed to contradict the connectionist account.

These data raise questions about the kinds of linguistic representations that were damaged in these patients and why damage yielded particular behavioral patterns. Ullman *et al.* (22) interpreted their results in terms of damage to different memory systems subserving rules and exceptions. In contrast, we view their results in terms of damage to the semantic and phonological codes that are characteristic of all words. In the past-tense generation task, a present-tense verb is presented auditorily or read aloud, and the subject says the past tense. Damage to phonology has a bigger impact on generating past tenses for novel verbs than for verbs with irregular pasts. A novel form, such as WUG, has no meaning, and so the only way to generate its past tense is by analogy to known phonological forms. Damage to phonology interferes with the capacity to do this. Damage to semantics, in contrast, has a bigger impact on generating irregular past tenses, such as TOOK. The fact that TAKE's past tense is TOOK is an idiosyncrasy of this verb. Producing the correct past tense for TAKE therefore requires identifying it as this specific word. If TAKE's semantic representation is damaged, information relevant to identifying it as this lexical item will be lost, and it will then pattern with phonologically similar verbs such as FAKE and TAME, which have rule-governed past tenses, yielding errors such as TAKED. Thus, we explain the observed deficits in terms of impairments to two types of lexical information, semantic and phonological, rather then memory systems organized around rules and exceptions.

One reason to pursue our approach is independent evidence that the patients in question have phonological or semantic deficits. Many Broca-type aphasics (24) and PD patients (28) have difficulty processing speech and/or planning articulatory output; these were the types of patients whose ability to generalize was impaired. Semantic impairments are characteristic of patients with PA (24) and AD (29); these patients were more impaired on irregular past tenses.

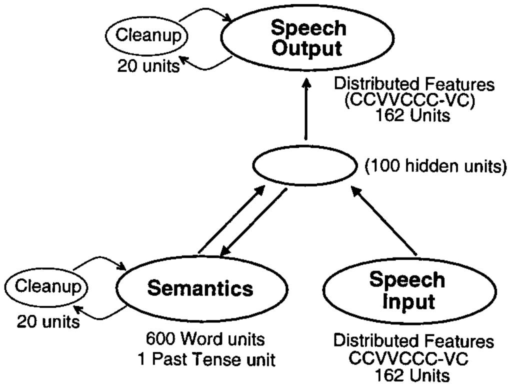

FIG. 1. Model architecture. Ellipses represent groups of units. Arrows represent connections between groups and the direction of information flow.

Simulation Model

We examined this account by implementing a connectionist model of some aspects of the past tense (Fig. 1). Words were represented in terms of the codes involved in the past-tense generation task: speech input (a representation of the sounds of words), speech output (a code produced in generating speech), and semantics. In humans, speech input consists of acoustic patterns, and speech output consists of sequences of articulatory gestures; as a simplifying assumption, we used the same distributed phonological representation for both codes. These representations employed a CCVVCCC-VC template (C, consonant; V, vowel), and each phoneme was represented in terms of 18 binary phonological featured.† Each word's vowel was aligned with the first V [VV was used for diphthongs such as /oy/ (BOY)]. Initial consonants were aligned with the C slots from right to left, and following consonants were aligned with C slots left to right. The final VC was used to represent the /Id/ syllable in words such as TASTED. All units in empty slots were set to 0.0.

The use of distributed representations allowed the model to represent degrees of phonological similarity between words; this property is relevant to the model's ability to generalize. Units on the speech output layer were connected to and from a set of "cleanup" units (30, 31). These units provided a way of representing nonlinearly separable phonological dependencies and made the computation of phonological output a dynamic process in which the model settled into a pattern over a series of time steps (32–34).

Each verb was represented by a unique node in the semantics layer. This localist representation does not capture semantic similarities between verbs; although this is crucial for other phenomena (35, 36), it is not important for the past tense. One additional unit was used to represent present/past tense semantics. The semantics layer was also connected to a cleanup layer of 20 units.

People acquire knowledge of language in the course of using it for different purposes. What is learned from a task, such as hearing, affects the ability to perform other tasks, such as speaking. We approximated this aspect of language by inter-

†Voiced, voiceless, consonantal, vocalic, obstruent, sonorant, lateral, continuant, noncontinuant, advanced tongue root (ATR), nasal, labial, coronal, anterior, high, distributed, dorsal, radical.

leaving training on four tasks. Speaking involved taking the semantic representation of a present or past tense verb as input and generating its phonology. Hearing involved taking the phonological code of a word as input and computing its meaning. Repeating involved taking a phonological code as input and generating the same code as output. Transforming involved taking the phonology of a verb and past-tense semantics as input and generating past-tense phonology. The model had to find a set of weights that allowed it to perform all of these tasks accurately.

The model was trained on the present and past tenses of 600 monosyllabic verbs, of which 64 had irregular past tenses. The present and past tenses of an additional 594 English verbs were used for the "repeating" trials, to give the model additional exposure to the structure of English phonology.‡ The probability of presenting each word during training was a function of its logarithmic frequency (37). Task probabilities were: speaking, 20%; hearing, 40%; repeating, 30%; and present–past transformation, 10%. Other simulations indicated that the model's performance was not highly sensitive to the exact proportions of trials of each type.

The network was trained by using the backpropagation-through-time algorithm (38). Each trial began with the random selection of an item and a task. The input appropriate for a given task was presented, and activation propagated throughout the network for seven time steps. Weights were adjusted based on the discrepancy between observed and expected patterns. Initial weights were randomized between −0.01 and 0.01; learning rate was: 0.005; a logistic activation function was used; error was calculated by using the cross-entropy measure (39).

Results

Training was halted after 2.6 million training trials, when performance was near asymptote. At this point, the network's accuracy was assessed on all words and tasks. Words were scored phoneme-by-phoneme, by using a Euclidean distance metric to determine the phoneme closest to the network's output; if the closest phoneme differed from the target phoneme, it was scored as incorrect. A computed word was scored as incorrect if the network produced any incorrect phonemes, including an empty slot instead of a phoneme and vice versa. Accuracy on the training set was: speaking, 99.8%; hearing, 99.5%; repeating, 98.2%; and transformation, 99.3%. Thus, the model had learned the training set quite accurately.

The model's capacity to generalize was assessed by using 20 nonce verbs used by Ullman *et al.* (22). The model was given the phonological code of the nonce verb and the past-tense semantics bit as input. The conjunction of these two types of information provided the basis for generating novel output. Using the same scoring criteria as above, the network generated correct past tenses on 90% of these items; on the two incorrect items, the network produced the past tenses of a similar known word (CLOGGED vs. CROGGED, SPURRED vs. SCURRED), a type of error that people also occasionally produce.

The model was then "lesioned" in two ways (Fig. 2). A phonological deficit was simulated by randomly severing connections between the speech output layer and its cleanup units. A semantic deficit was simulated by severing a proportion of connections between the semantics layer and its cleanup units, and Gaussian noise was added by multiplying each computed output in the semantics layer by a different random value (mean = 1.0) at each time step; degree of Gaussian noise was

‡The model's performance depends in part on its capacity to represent English phonology adequately. A child acquires this knowledge in the course of learning a language, not merely learning present- and past-tense verbs. Training on these additional words provided additional exposure to English phonology. Other types of words (e.g., nouns) could also be used for this purpose.

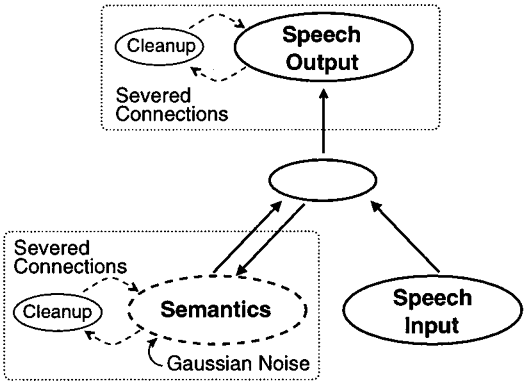

FIG. 2. Phonological (*Lower*) and semantic (*Upper*) lesions.

manipulated by changing its SD. The effects of these lesions on the transformation task were assessed by using the 20 regular, 16 irregular and, 20 nonword verbs from Ullman *et al.* (22).

Figs. 3 and 4 present data from simulations in which the amount of each type of damage was parametrically varied. Each data point is the mean of 10 simulations that differed in terms of which connections were randomly lesioned. Damage to phonology affected performance on all three types of verbs, but the effect was largest for nonwords. Damage to semantics also affected all three types of verbs, but here the effect was largest for irregular verbs. These results indicate that the "double dissociation" observed across patient groups can be replicated by introducing different types of damage in a system that does not include separate rule and exception mechanisms.

The model's errors under different types of damage were also broadly consistent with the patient data. Semantic damage created regularizations of irregular past tenses, such as KEEPED; phonological damage did not. With phonological damage, the model produced the same types of errors as patients, including irregularization errors (e.g., BAKE–BOKE, omissions of the past-tense inflection (e.g., BAKE–BAKE), and other phonological deviations (e.g., VASK–BASKED). Semantic damage produced fewer of these errors.

Other Phenomena

Other aspects of Ullman *et al.*'s (22) data need to be considered with caution. Data were reported for a subset of the AD (5/24) and PD (5/28) patients who were tested. The net number of patients per group was small (either 5 or 1), making it difficult

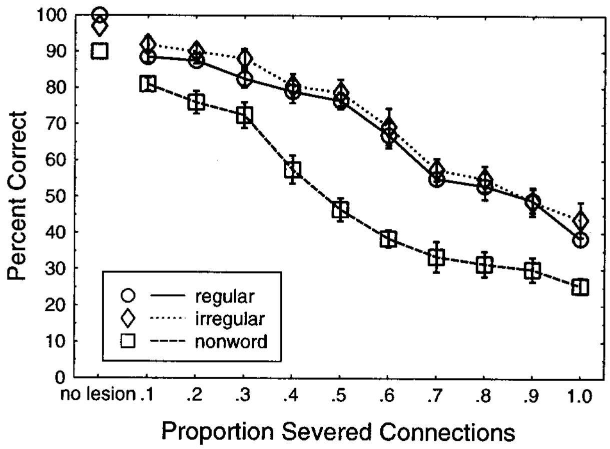

FIG. 3. Effects of phonological damage on models' ability to produce three types of English past tense verbs, averaged over 10 simulated lesions per degree of severity. Severity was manipulated by increasing the proportion of randomly severed connections to and from the Speech Output's cleanup layer (10%–100%).

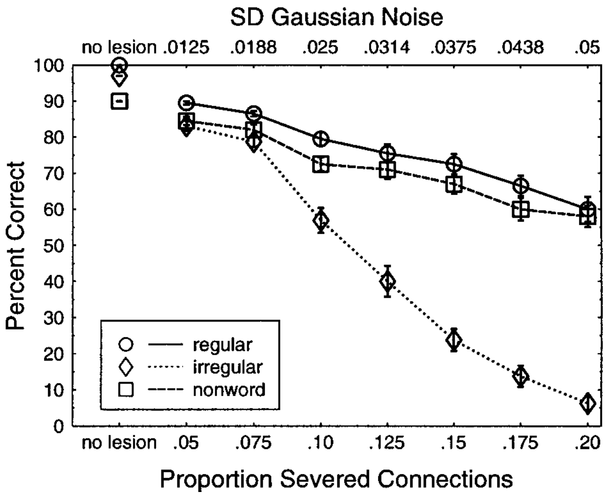

FIG. 4. Effects of semantic damage on models' ability to produce three types of English past-tense verbs, averaged over 10 simulated lesions per degree of severity. Severity was manipulated by increasing the proportion of randomly severed connections to and from the Semantics cleanup layer (5–20%) and the SD of Gaussian noise added to activation in the Semantics layer (0.0125–0.05).

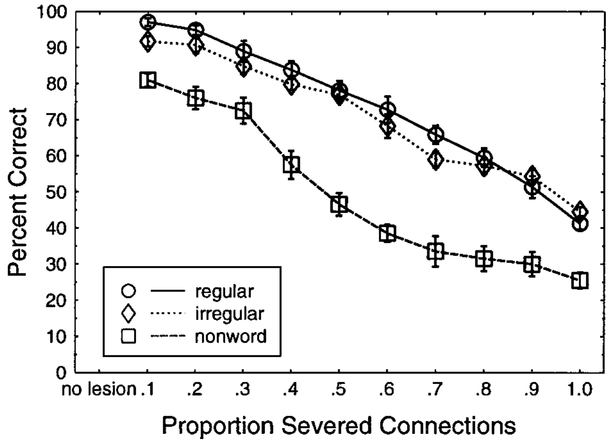

FIG. 6. Effects of phonological damage on models' ability to produce the past tenses of high-frequency English past-tense verbs and nonwords averaged over 10 simulated lesions per degree of severity. The regular and irregular items were more closely matched in terms of frequency than were the items in Ullman *et al.* (22) study. Severity was manipulated by increasing the proportion of randomly severed connections to and from the speech output's cleanup layer (10–100% severed connections.)

to assess measurement error and effect reliability. (It is noteworthy that the one group with a larger n, the 17 Huntington's Disease patients, did not show reliable differences between types of verbs.) Some of the patients could not perform the past-tense generation task and were tested instead on a different task (reading the verbs aloud). The number of items per condition was small, and 4 of the original 20 irregular items were excluded from the analyses post hoc. The regular and irregular stimuli also were not closely matched in terms of two factors that affect normal performance, frequency and phonological complexity. The regular verbs were lower in frequency than the irregulars [mean frequencies from ref. 37: regular, 33; irregular, 140; $t(34) = 2.30, P < 0.05$]. The regular past tenses were also more phonologically complex, containing more consonant clusters in word-initial and -final positions. These aspects of the study make us wary of attempting to closely simulate data for individual patients or patient groups. However, we can show that the kinds of data that Ullman *et al.* (22) reported are consistent with our model and our account of the bases of impairments in verb morphology.

Phonological Deficits. Fig. 5 presents means from the five PD patients reported in Ullman *et al.* (22), along with the results of simulations in which 30% of the phonological clean-up connections were randomly severed. The model data are the means of 10 simulations. Like the patients, the phonologically lesioned models were impaired on all three types of stimuli, with nonwords more impaired than regulars, $t(9) = 3.35, P < 0.01$, and irregulars, $t(9) = 4.26, P < 0.01$, which did not differ reliably from each other, $t(9) = 1.80, P > 0.05$. The lesioned models were less impaired on the task of repeating the target past tenses (regular, 90%, irregular, 99%, nonword, 91%). Thus, the deficits in generating past tenses were not merely due to difficulties with the target words themselves. Rather, the effects of phonological damage are largest for tasks that place heavy demands on phonological processing, including nonword reading and past tense production (40).

The claim that the PD group suffered from a deficit to the rule module rests on the observation that their mean performance on both types of rule-governed items (regulars and nonwords) was lower than on irregulars. However, as Fig. 5 indicates, the difference between regulars and irregulars was not statistically reliable. Moreover, this small difference appears to be related to the fact that the regular verbs in the Ullman *et al.* study (22) were significantly lower in frequency than the irregulars. We also assessed the phonologically damaged models using a set of 58 regular and 58 irregular verbs that were more closely equated in frequency (we achieved this by including all irregular items in the training set that could be matched with a regular item of a similar frequency). These items yielded a small advantage for regulars over irregulars (Fig. 6). Finally, we have observed, using Ullman *et al.*'s stimuli, that introducing random phonological damage occasionally produces a small advantage for irregulars over regu-

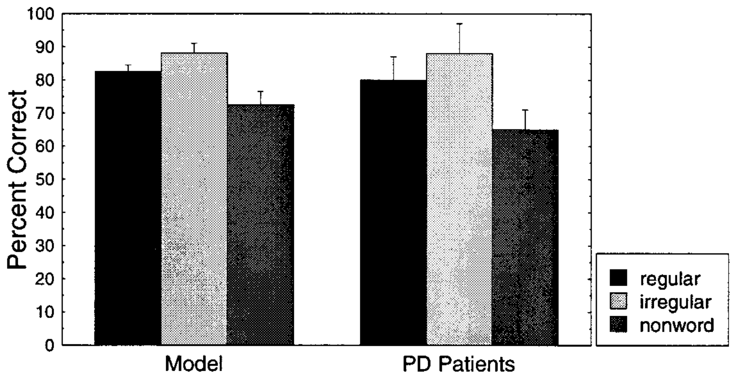

FIG. 5. Effects of phonological damage on production of regular, irregular, and nonword past tenses. Model data are means of 10 simulations in which phonology was damaged (30% severed connections); patient data are means for five PD patients reported by Ullman *et al.* (22).

Table 1. Results for individual lesioned models demonstrating worse performance on regular and nonwords compared to irregular words.

Verb Type		
Irregular	Regular	Nonword
63	35	20
63	35	30
50	25	20
56	30	35
63	30	20
56	30	25
56	25	30
56	45	30
50	20	20

Instances are taken from 50 random lesions. Data are given as % correct.

lars. In short, in both patients and models, phonological impairment had a large impact on nonword generalization, with less effect on both regular and irregular verbs.

Ullman *et al.*'s (22) final bit of data concerning impaired generalization is from a single anterior aphasic patient, FCL, whose behavior was unusual. Whereas other patients with phonological deficits performed similarly on regular and irregular verbs, patient FCL scored 69% correct on irregulars, but 20% and 5% correct on regulars and nonwords, respectively, perhaps the strongest evidence for an impaired "rule module." We investigated whether the model could yield a pattern similar to FCL's by creating a severe phonological lesion in the model (severing all connections to and from the speech cleanup layer and adding Gaussian noise with a SD of 0.25). Fifty models randomly lesioned in this way were tested on verbs matched to the ones used in testing FCL. The simulations yielded a distribution of impairments, including many exhibiting a severe impairment on regulars and nonwords compared to irregulars (Table 1).

Patients with such extreme dissociations are perhaps the most challenging for our account, and so it will be important investigate them further. One direction for future research concerns other perceptual factors that affect the processing of regular past-tense morphology. In English, the past tense is formed by adding phonemes lacking perceptual salience to the ends of words (/t/, /d/, or /Id/), a factor that Leonard *et al.* (41) have shown is related to the errors produced by children with developmental language impairments. Hoeffner and McClelland (42) developed a connectionist model similar to the one presented here that showed how a general phonological impairment could affect the processing of this less-salient material, thereby affecting the network's performance on verbs with regular endings. Our model does not address the

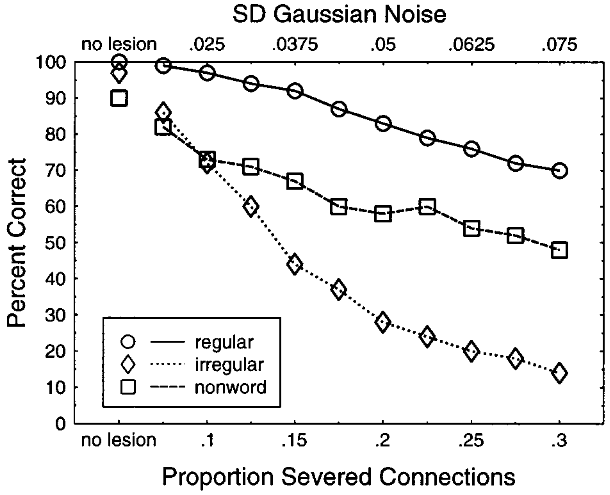

FIG. 8. Effects of semantic damage on models' ability to produce the past tenses of high-frequency verbs and nonwords, averaged over 10 simulated lesions per degree of severity. Severity was manipulated by increasing the proportion of randomly severed connections to and from the semantics cleanup layer, along with the SD of Gaussian noise added to activation in the semantics layer (5–20% severed connections, 0.0125–0.075 SD noise).

perceptual salience issue; however, Hoeffner and McClelland's work suggests that introducing this factor could magnify the dissociation between regular and irregular forms.

Semantic Deficits. Ullman *et al.* (22) presented data from a group of five AD patients, a group of five PA patients, and JLU, a single PA patient. These data (Fig. 7) demonstrate a small impairment on nonwords compared to regular words and a much larger impairment on irregular words. For comparison, we present the effects of a semantic impairment on the model's performance. The lesion involved severing 10% of cleanup connections and adding 0.025 SD noise. This produced a pattern very similar to both the AD and PA patient groups and patient JLU, with performance on irregulars significantly worse than on both nonwords, $t(9) = 3.10$, $P < 0.025$, and regulars, $t(9) = 6.27$, $P < 0.001$.

Ullman *et al.*'s (22) interpretation that the AD and PA groups suffer from a deficit to an exception module rests on the observation that regulars and nonwords patterned together, with performance on both significantly better than on irregulars. Again, however, it is necessary to consider the extent to which this pattern was caused by the fact that the regular verbs used in testing the patients happened to be relatively low in frequency. We retested the semantically-impaired models on the 58 regular and 58 irregular verbs discussed above. As Fig.

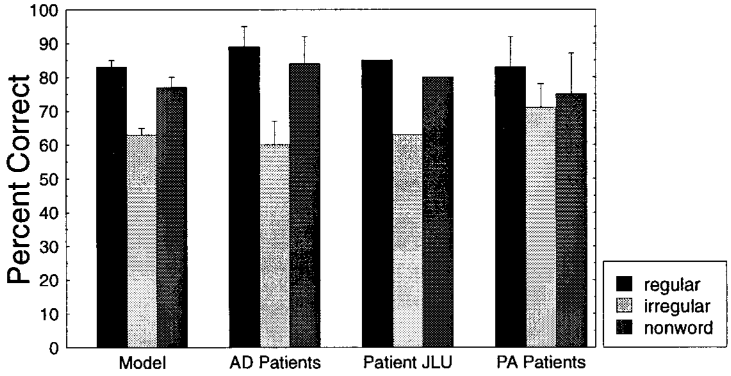

FIG. 7. Effects of semantic damage on past tense production. Model data are means of 10 simulations in which semantic representations were damaged (10% severed connections, 0.025 SD noise); patient data are means from five AD patients, five PA patients, and PA patient JLU. Patient data are from Ullman *et al.* (22).

8 indicates, with these matched items regular words and nonwords no longer pattern together. In the model, semantic damage impairs the generation of past tenses for both regular words and nonwords because it affects the use of the past tense bit. The effect is larger for nonwords because the model has not seen them before, and still larger for irregular verbs. In summary, semantic damage has a large effect on generating irregular past tenses and smaller effects on regular words and nonwords.

Discussion

The purpose of the present research was to explore a hypothesis concerning the bases of impairments in verb morphology suggested by connectionist accounts of language. The finding that some patients are more impaired in generating irregular past tenses and others in generating novel past tenses intuitively seems to implicate separate brain mechanisms dedicated to these functions. The validity of this type of interpretation of double-dissociation data has been strongly called into question by analyses of such effects in connectionist modeling (43) and by studies of neurocognitive disorders (44). Our simulations suggest that the behavioral deficits arise from impairments in two types of lexical information, semantics and phonology, which happen to make somewhat different contributions to generating irregular forms and generalization. The fact that brain injury can affect the use of phonology or semantics is well established from studies of other aspects of language. Our analysis is also consistent with independent evidence that AL and PD are associated with phonological impairments and that AD and PA are associated with semantic deficits. The simulations demonstrate that damage to these types of information does differentially affect performance on irregulars and nonwords.

Ullman *et al.*'s theory, in contrast, assumes the traditional distinction between rule-governed forms and exceptions. In their view, one mechanism is used to generate the past tense of BAKE and a completely different one for TAKE. These mechanisms are thought to have different computational characteristics: one involves applying rules, the other using associations; one is said to be "procedural," the other "declarative." Our theory is different: there is a single, distributed network that represents people's knowledge of words. The same network structure is used in processing all words; in doing so, the network captures partial regularities among forms that the other theory treats as categorically different. The explanation for the behavioral dissociations lies with the fact that the network includes distinct phonological and semantic representations, which have different realizations in the brain, can be differentially affected by neuropathology and have different effects on known words and generalization.

More detailed aspects of these phenomena await further investigation. On the behavioral side, it will be important to examine a larger number of patients in close detail, assessing their knowledge of both the past tense and other aspects of phonology and semantics. Without this information, it is difficult to justify developing models that provide close quantitative fits to the data. On the modeling side, it would be preferable to study impairments after brain injury in the context of a model that accounts for other aspects of acquisition and normal performance. Combining these behavioral and computational approaches is likely to greatly extend our knowledge of the past tense and its brain bases.

We thank our colleagues James L. McClelland, Karalyn Patterson, Matt Lambon-Ralph, and David Plaut for very helpful comments on an earlier draft of this paper. This research was supported by National Institute of Mental Health Grants PO1-MH47566 and RO1-MH58723. M. J. was supported by an National Sciences and Engineering Research Council (Canada) postgraduate fellowship (PGS-B 214574) and M. S. by an National Institute of Mental Health Research Scientist Development Award (KO2-MH01188).

1. Chomsky, N. (1965) *Aspects of the Theory of Syntax* (MIT Press, Cambridge, MA).
2. Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin and Use* (Praeger, New York).
3. Chomsky, N. (1995) *The Minimalist Program* (MIT Press, Cambridge, MA).
4. Rumelhart, D. E. & McClelland, J. L. (1986) in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, eds. McClelland, J. & Rumelhart, D. (MIT Press, Cambridge, MA, Vol. 2).
5. Elman, J. L., Bates, E. A, Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking Innateness: A Connectionist Perspective on Development* (MIT Press, Cambridge, MA).
6. Seidenberg, M. S. (1997) *Science* **275,** 1599–1603.
7. Roeper, T. & Williams, E., eds. (1987) *Parameter Setting* (Kluwer, Boston).
8. Pinker, S. & Prince, A. (1988) *Cognition* **28,** 73–193.
9. Pinker, S. (1991) *Science* **253,** 530–535.
10. Berko, J. (1958) *Word* **14,** 150–177.
11. Seidenberg, M. S. & McClelland, J. L. (1989) *Psychol. Rev.* **96,** 523–568.
12. Bauer, L. (1983) *English Word-Formation* (Cambridge Univ. Press, Cambridge, U.K.).
13. Halle, M. & Mohanan, K. (1985) *Linguistic Inquiry* **16,** 57–116.
14. Stemberger, J. P. & MacWhinney, B. (1986) *Cognit. Psychol.* **18,** 329–354.
15. Chomsky, N. & Halle, M. (1968) *The Sound Pattern of English* (MIT Press, Cambridge, MA).
16. Plunkett, K. & Marchman, V. (1991) *Cognition* **38,** 43–102.
17. Daugherty, K. & Seidenberg, M. S. (1992) in *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (Lawrence Erlbaum, Hillsdale, NJ), pp. 259–264.
18. MacWhinney, B. & Leinbach, J. (1991) *Cognition* **29,** 121–157.
19. Seidenberg, M. S. & Hoeffner, J. H. (1998) *Language* **74,** 104–122.
20. Jaeger, J. J., Lockwood, A. H., Kemmerer, D. L., Van Valin R., Jr., Murphy, B. W. & Khalak, H. G. (1996) *Language* **72,** 451–497.
21. Marslen-Wilson, W. & Tyler, L. (1997) *Nature* (*London*) **387,** 592–594.
22. Ullman, M. T., Corkin, S., Coppola, M., Hicock, G., Growdon, J. H., Koroshetz, W. J. & Pinker, S. (1997) *J. Cognit. Neurosci.* **9,** 266–276.
23. Shallice, T. (1988) *From Neuropsychology to Mental Structure* (Cambridge Univ. Press, Cambridge, MA).
24. Caplan, D. (1992) *Language: Structure, Processing and Disorders* (MIT Press, Cambridge, MA).
25. Zigmond, M. J. & Stricker, E. M. (1984) *Life Sci.* **35,** 5–18.
26. Brun, A. & Englund, E. (1981) *Histopathology* **5,** 549–564.
27. Kemper, T. (1994) in *Clinical Neurology of Aging*, eds. Albert, M. & Knoefel, J. (Oxford Univ. Press, New York), pp. 3–67.
28. Grossman, M., Carvell, S., Stern, M. B., Gollomp, S. & Hurtig, H. I. (1992) *Brain Lang.* **42,** 347–384.
29. Hodges, J. & Patterson, K. (1995) *Neuropsychologia* **33,** 441–459.
30. Hinton, G. E. & Shallice, T. (1991) *Psychol. Rev.* **98,** 74–95.
31. Plaut, D. C. & Shallice, T. (1993) *Cognit. Neuropsychol.* **10,** 377–500.
32. McClelland, J. & Rumelhart, D. (1981) *Psychol. Rev.* **88,** 375–407.
33. Plaut, D. C., McClelland, J. L., Seidenberg, M. & Patterson, K. E. (1996) *Psychol. Rev.* **103,** 56–115.
34. Harm, M. W. & Seidenberg, M. S. (1999) *Psychol. Rev.*, in press.
35. Allen, J. (1997) in *Proceedings of the GALA '97 Conference on Language Acquisition*, eds. Sorace, A., Heycock, C. & Shillcock, R. (Univ. of Edinburgh, Edinburgh), pp. 300–305.
36. Gleitman, L. (1990) *Lang. Acquisition* **1,** 3–55.
37. Francis, W. N. & Kuçera, H. (1982) *Frequency Analysis of English Usage* (Houghton–Mifflin, Boston).
38. Williams, R. J. & Peng, J. (1990) *Neural Comput.* **2,** 490–501.
39. Hinton, G. E. (1989) *Artif. Intell.* **40,** 185–234.
40. Patterson, K., Suzuki, T. & Wydell, T. N. (1996) *Cognit. Neuropsychol.* **13,** 803–822.
41. Leonard, L., Bortolini, U., Caselli, M., McGregor, M. & Sabbadini, L. (1992) *Lang. Acquisition* **2,** 151–179.
42. Hoeffner, J. J. & McClelland, J. L. (1993) in *Proceedings of the 25th Annual Child Language Research Forum*, ed. Clark, E. (Centre for the Study of Language and Information, Stanford, CA), Vol. 25, pp. 38–49.
43. Plaut, D. C. (1995) *J. Clin. Exp. Neuropsychol.* **17,** 291–321.
44. Farah, M. J. (1994) *Behav. Brain Sci.* **17,** 43–104.