

Are words represented by nodes?

GREGORY O. STONE and GUY C. VAN ORDEN
Arizona State University, Tempe, Arizona

Semantic priming in a lexical decision task was investigated with concurrent pretarget and posttarget primes. The posttarget prime also served as a pattern mask of the lexical decision target. Forward priming is defined as the presence of a semantically related pretarget prime and an unrelated posttarget prime. Backward priming is defined as the presence of a semantically related posttarget prime and an unrelated pretarget prime. Forward and backward priming were compared both when the nonword foils were "legal" and when they were "illegal" with respect to English orthography. Predictions were derived for two general approaches to word recognition: spreading-activation and expectancy-set theories. Both approaches assume that word representations occupy distinct, nonoverlapping locations in memory. Backward-prime facilitation was equivalent to forward-prime facilitation when nonword foils were illegal; however, backward-prime facilitation was not significant when nonword foils were legal. These results challenge both approaches. The proposed solution uses semantic-space (distributed) representations that are *functionally* unitized by a resonant matching (verification) process.

Spreading-activation and expectancy-set theories of word recognition make specific predictions about the time course of semantic priming. Despite the differences between these approaches, the predictions of both derive, in part, from their shared assumption that lexical representations are structurally unitized (i.e., each lexical representation is a separate, distinct structure). In effect, both approaches assume that words are represented by nodes in the broad sense of structurally distinct representations. In this article, we report the investigation of these predictions in a bidirectional priming, lexical decision paradigm. The facilitation due to a forward (pretarget) prime is compared with that due to a backward (posttarget) prime. This is an extreme version of varying prime-target onset asynchrony. We first introduce the details of this bidirectional priming technique and then examine spreading-activation and expectancy-set theories, deriving for each predictions concerning patterns of forward and backward priming.

The Bidirectional Priming Paradigm

In the bidirectional priming paradigm, each trial begins with the presentation of a pretarget priming word, which can be either semantically related to the subsequent lexi-

cal decision target or unrelated to it. This prime remains visible long enough to be clearly recognized. A lexical decision target item (a word or a nonword) is then presented briefly in the same location. Upon offset, it is replaced, again in the same location, by a posttarget priming word, which can also be either semantically related to the target or unrelated to it. Subjects make a lexical decision judgment (i.e., respond "word" or "nonword") about the target. The pretarget prime also serves as a fixation stimulus, indicating where the target will occur, and the posttarget prime also serves as a backward-pattern mask. The dependent variable is the proportion of errors on word trials (i.e., responses of "nonword" to word targets).

The principal manipulation is the semantic relatedness of the primes and the target. In the forward-prime condition, the pretarget prime is semantically related to the target and the posttarget prime is unrelated to it. In the backward-prime condition, the posttarget prime is related to the target and the pretarget prime is unrelated to it. In the unrelated-prime condition, which serves as a baseline for measuring facilitation in both related-prime conditions, neither prime is related to the target.

Finally, the proportion of related-prime trials is reduced by the inclusion of unrelated prime-filler trials to discourage controlled priming by the pretarget prime (Tweedy, Lapinski, & Schvaneveldt, 1977). Such controlled priming is a concern because controlled inhibition from the unrelated pretarget prime could partially counteract facilitation from the related posttarget prime in the backward-prime condition. Thus, controlled forward priming could produce an underestimate of backward-priming effectiveness.

Although several investigators have reported facilitation of word recognition due to semantically related word masks (Jacobson, 1973, 1974; Jacobson & Rhinelanders,

This research partially fulfilled the requirements for a doctoral degree for Greg Stone at the University of California, San Diego. His work on this research was supported in part by an intramural grant from the University of California, San Diego to Jeffrey Miller and a grant from the Office of Naval Research (ONR N00014-83-K0337) to Stephen Grossberg. We wish to thank Greg Stone's dissertation committee (Jeff Miller, Jeff Ellman, Don MacCleod, Jay McClelland, Dave Rumelhart, and Tony Sebold) for their invaluable guidance in the development of this research. We also wish to thank Ron Epperlein, Don Homa, Peter Killeen, and Ennio Mingolla for helpful comments on earlier drafts, and Bancroft Dutton, Steve Grossberg, George Mandler, Christine Stone, Cindy Suchta, and Ruby Woods for various other forms of assistance. Reprint requests should be sent to Greg Stone, Department of Psychology, Arizona State University, Tempe, AZ 85287.

1978; Taylor & Chabot, 1978), these studies contained a serious methodological flaw. Target-mask pairs were repeated at increasing interstimulus intervals until subjects could correctly identify the target word. The interstimulus interval at which subjects identified the target was taken as a measure of mask effectiveness. Because the mask (backward prime) on one trial could prime recognition of the same target on the next trial, the decrease in mask effectiveness (priming) for semantically related word masks could have resulted from facilitation due to forward priming by the mask from the preceding trial.

Backward priming without masking was investigated by Kiger and Glass (1983), who presented primes in a location adjacent to that of a lexical decision target. Reaction times to the target were the dependent measure. They found significant prime facilitation of target recognition when prime onset followed target onset. Although Kiger and Glass argued that this result challenges expectancy-set theories, it is possible that the appearance of a prime after target presentation draws attention from processing of the target, allowing recognition of the backward prime before processing of the target. Although this would slow lexical decision times, their study did not provide the baseline needed to dismiss this alternative interpretation. It is interesting to note that Kiger and Glass argued that a backward prime that also serves as a pattern mask should not facilitate target processing.

The bidirectional priming paradigm addresses this argument while avoiding the problems that arose in previous studies of backward priming. Lexical decision judgments allow a single presentation of a target for each subject. Moreover, unlike previous backward-priming studies, the bidirectional priming paradigm allows comparison of facilitation by forward and backward primes under otherwise identical presentation conditions, in particular, conditions in which a posttarget context always disrupts target processing in a manner that places a premium on rapid decisions concerning the target.

We now consider spreading-activation and expectancy-set theories, deriving predictions from each concerning patterns of forward and backward priming in this task.

Spreading-Activation Theories

In network theories of semantic processing, each semantic concept is represented by a node, with links between nodes for related concepts (Collins & Loftus, 1975; Collins & Quillian, 1969). The most natural means of access to a semantic network during word recognition is provided by spreading-activation theories, which assume that words are represented by nodes and that recognition of a word causes a spread of activation to nodes for related words (Morton, 1969). This spread of activation facilitates recognition of words occurring in a semantically related context. Because associative links join spatially separated nodes, spreading activation should occur with some nonzero time course.

Ratcliff and McKoon (1981) tested for a nonzero time course of spreading activation. Their subjects first studied

test paragraphs and then (in a subsequent recognition task) were asked to judge whether a test word had occurred in one of the study paragraphs. Each recognition-test target word was preceded by a prime that had appeared either in the same study paragraph or in a different study paragraph. Primes from the same-study paragraph had appeared either near or far from the target word in that paragraph. They found greater facilitation for primes that had previously appeared near the target in the study paragraph, but found no change in facilitation as prime-target onset asynchrony in the recognition task decreased. (Prime-target onset asynchrony in the recognition task was varied from 50 to 350 msec). This result suggests that spreading activation would need to be very rapid if it occurs during the recognition of a previous occurrence in a text.

The results of Ratcliff and McKoon (1981) do not, however, necessarily address access to long-term semantic memory. Distances between nodes were defined in terms of the distance between words in the viewed text. Considering that pre-experimental relatedness does not affect priming in this paradigm (McKoon & Ratcliff, 1980), text processing could be working with episodic representations that are distinct from semantic representations (Carroll & Kirsner, 1982; Neely & Durgunoglu, 1985; Shoben, Wescourt, & Smith, 1978; however, cf. McCloskey & Santee, 1981).

Unlike the priming paradigm of McKoon and Ratcliff (1980; Ratcliff & McKoon, 1981), the bidirectional priming paradigm examines the time course of facilitation during access to long-term semantic memory. With the proper controls, the comparison of forward and backward priming should optimize the likelihood of detecting a nonzero time course for the spread of activation between semantically related nodes. The central argument is that as prime-target asynchrony decreases, it becomes more probable that a lexical entry will be selected before spreading activation from the prime has reached full efficacy.

At least two mechanisms for spreading activation can be distinguished in terms of the time course of spreading activation. First, spreading activation may build up over time at the receiving node. This assumption is natural in a system in which activation can begin to spread before the prime's node exceeds threshold. The relatively long duration of the forward prime would allow spreading activation (facilitation) to come asymptotically close to its maximum value. In addition, facilitation from a backward prime would be expected to influence target processing, but not to reach full effectiveness by the time a lexical decision was made for the target. Backward priming should occur but it should provide less facilitation than forward priming.

Second, activation can arrive at the receiving node in a discrete pulse. In this case, trials in which activation from a related prime arrives at the target word's node before a lexical decision is made will exhibit full facilitation. Trials in which activation does not arrive before a lexical decision is made will exhibit no facilitation. Again,

the long duration of the forward prime makes it likely that such a pulse of activation will arrive before a lexical decision is made. On the other hand, because of the strong time constraints on backward-prime facilitation, it is likely to comprise a mixture of trials with effective and ineffective spread of activation, so that backward priming will be much less effective than forward priming.

In either case, the most natural prediction for spreading-activation theories is that backward priming will facilitate lexical decisions for related targets but that this facilitation will not be as great as that provided by forward priming, in which spreading activation can reach full effectiveness. Moreover, if activation does not begin to spread from the prime's node until it exceeds threshold, one would expect little or no backward priming.

Expectancy-Set Theories

Like network theories, schema theories of semantic memory assume structural unitization of concepts, or nodes, in the broad sense. Unlike the nodes in network theories, however, schemata have internal structure. Semantic relationships between concepts are reflected in similar components of the internal structure of their schemata. Thus, unlike the links of network theories, semantic relationships are not structurally manifest in the organization of semantic memory.

The most natural means of access to a schema-based semantic memory is provided by expectancy-set theories. Despite differences in their assumptions about early stages of processing (cf. verification theories: Becker, 1976, 1979, 1980; Becker & Killion, 1977; Paap, Newsome, McDonald, & Schvaneveldt, 1982, vs. bin theories: Forster, 1976), all expectancy-set theories of lexical selection assume that a candidate set of lexical schemata, generated by the earlier stages of processing, is used in a serial-matching process in the final (verification) stage of processing. The candidate schemata are matched against the perceptual schema for the target word one at a time, with the schemata for more familiar words being matched before those for less familiar words. This matching process continues until the match between a memory schema and the perceptual schema exceeds a match criterion. In all expectancy-set theories, semantic priming effects are restricted to this final stage of lexical processing.

When a word appears in context, a candidate set of lexical schemata for words consistent with that context can be generated before earlier processing stages have been completed. The matching thus begins earlier than it does when no context is present. Schemata from this "semantic" search set are matched against the perceptual schema for the target first. If no match is found, the bottom-up, or visual, search set is searched as before. As a result, recognition is faster for words whose schemata are contained in this semantic search set (i.e., words related to the context) than for words whose schemata are not contained in the semantic search set (i.e., words unrelated to the context).

The predictions of expectancy-set theories for the bidirectional priming paradigm depend on assumptions about backward-pattern masking. However, assumptions about masking are not an intrinsic part of expectancy-set theory. Only in Paap et al. (1982) is an account of masking made explicit. They argued that a backward-pattern mask prevents or terminates the verification stage. Given this assumption, the semantic content of the mask should have no effect on processing of the target, since initiation of mask processing terminates the stage of target processing responsible for facilitation. Furthermore, since the forward prime also facilitates recognition of the target in the verification stage, one would also expect little or no forward priming.

A viable alternative is that masking disrupts, but does not terminate, the verification process. In this case, verification matching is performed using a decaying perceptual trace (Sperling, 1960). Schemata matched sooner in the serial search are verified against a more complete perceptual schema. Thus, words whose schemata are matched early in the search (e.g., words related to the context) will yield fewer errors than will words whose schemata are matched later (e.g., words unrelated to the context). Forward priming should occur, but backward priming can occur only if verification matching is delayed until the mask word is recognized and a semantic search set is generated for it. Considering the time necessary for this process, and the rapid decay of perceptual information, the most natural prediction in this case is little or no backward priming.

Summary of Predictions

The "most natural" predictions for spreading-activation and expectancy-set theories can be grouped according to the possible outcomes of the bidirectional priming experiment.

First, no facilitation in either the forward- or the backward-priming conditions would support the assumption that masking prevents verification matching.

Second, facilitation due to forward priming, but not due to backward priming, would be consistent with a matching process that uses a decaying perceptual trace under masking conditions and with activation that spreads very slowly or cannot begin spreading until a prime exceeds threshold.

Third, facilitation due to forward and backward priming, with significantly more facilitation in the forward-priming condition, would be consistent with a rapid spread of activation.

In the present experiment, we introduce the bidirectional priming paradigm and test these predictions. Furthermore, a manipulation of nonword lexicality is included. The term "lexicality" applied to nonwords concerns the degree to which a nonword is similar to words. "Illegal" nonwords violate the rules of English orthography and are visually similar to few, if any, words. "Legal" nonwords follow the rules of English orthography and are, in general, visually similar to a number of words.

The manipulation of nonword lexicality (in this case legal vs. illegal) is of interest because it has been suggested that lexical decision judgments can be made on the basis of the output from several distinct subsystems. In particular, Shulman and Davison (1977) argued that rapid decisions about pronounceability could be used to make a lexical decision before processing in the lexicon had finished. This argument was based on their finding of reduced semantic priming with illegal nonword foils, relative to that with legal nonword foils. A key assumption was that the semantic context only affects processing in the lexicon. Effects of nonword lexicality have also been demonstrated in several other studies (James, 1975; Shulman, Hornak, & Sanders, 1978). Inclusion of the nonword-lexicality manipulation provided a further test of the argument that strategic control in word recognition is due to shifting reliance on lexical and nonlexical sources of information.

Bidirectional priming was first investigated in the first two experiments of Stone (1985). Several methodological issues remained, however, which could not be resolved with a single choice of methods. This experiment (Experiment 3 of Stone, 1985) differed from Experiments 1 and 2 of Stone (1985) on these points as a check against the possibility that the bidirectional priming results were due to a methodological artifact. The important differences from Stone's Experiments 1 and 2 and their significance will be noted in the Method section. Furthermore, because this experiment was an attempt at replication and the effects of interest were not directly reflected in the sources of the omnibus ANOVA, a set of planned comparisons was designed as the principal analysis.

METHOD

Subjects

Forty-eight undergraduates at the University of California, San Diego served as subjects in two sessions on 2 consecutive days. The subjects were paid \$6 per session.

Apparatus

A Gerbrands three-channel tachistoscope was used to present the stimuli, which were printed in the center of 4×6 index cards. The targets were printed in lower-case IBM prestige elite 12-pitch type and the context words were printed in upper-case IBM orator 10-pitch type. The apparent distance of the stimuli was 80 cm, so that a five-letter target subtended a visual angle of 46.2 min horizontal × 12.9 min vertical, whereas a five-letter context word subtended a visual angle of 55.4 min horizontal × 17.2 min vertical. The background field subtended a visual angle of 9.2° horizontal × 6.4° vertical.

Experiments 1 and 2 of Stone (1985) employed binocular masking with luminous stimuli on a dark background. To ensure that the bidirectional priming effects occur for central, interruptive masking, polarizing filters of opposite orientation were used to allow dichoptic presentation of the target and the mask (Turvey, 1973). The pretarget context word was presented binocularly, allowing the subjects to establish binocular convergence. The target was then presented to the left eye and the posttarget context to the right eye, producing dichoptic pattern masking of the target by the posttarget context. The perceived brightness of each channel was roughly equated. Responses were verbal and were recorded manually by the experimenter. Experiments 1 and 2 of Stone (1985) were

computer-controlled and the experimenter was not present during the experiment. Thus, results constituting a successful replication are unlikely to be due to experimenter bias.

Stimuli

Forty-eight semantically related word pairs were generated on the basis of a high-associativity rating in Palermo and Jenkins (1964), typicality of category membership based on category norms (Battig & Montague, 1969), and the experimenter's impression of semantic relatedness (Fischler, 1977). Experiments 1 and 2 of Stone (1985) used a different stimulus set (based on Palermo & Jenkins's, 1964, association norms). Replication with a new stimulus set was a concern, not only for the usual reasons, but also because mask effectiveness depends upon complex interactions between the visual structure of targets and masks (McClelland, 1978; Smith, Haviland, Reder, Brownell, & Adams, 1976). Due to the nature of the bidirectional priming paradigm, it is impossible to fully control for such differences between the forward- and backward-priming conditions, since the mask must be different. The replication thus reduces the likelihood that the bidirectional priming results are due to such a difference.

If one member of a related pair was longer than the other, it was always assigned to the role of prime. This, and the larger type used for the primes, guaranteed that the target always appeared within the perimeter of the primes.

A second, unrelated prime, which matched the related prime for length, word frequency (Kučera & Francis, 1967), part of speech (Morris, 1969), and, whenever possible, number of syllables was generated for each target. In addition, each target was paired with another target of the same length such that their context words were matched for length, word frequency, and, whenever possible, part of speech and syllable length. Thus, unrelated primes could be provided for each target by exchanging prime pairs between yoked targets.

Two nonwords were generated from each target. The legal nonword was created by changing either the first or the last letter of the target word so as to roughly equate the bigram frequencies of the word and the nonword. Pseudohomophones of the source word (e.g., DOCTOR nerse) were avoided. Furthermore, whenever possible, the envelope of the nonword was different from its source word (e.g., letters with ascenders in lower case were changed to letters with descenders or without either, etc.). Thus, the legal nonwords were designed to be as discriminable as possible from their source words, within the constraint of changing a single letter. Experiments 1 and 2 of Stone (1985) avoided such pseudoprimes (e.g., DOCTOR gurse) since pseudoprimes confounds with the legal/illegal manipulation. Inclusion of pseudoprimes in this experiment served as a test for sensitivity of bidirectional priming to pseudoprimes. The illegal nonword was created by taking an anagram of the target word. On nonword trials, the appropriate nonword was substituted for the word target from which it was generated. The experimental stimuli are presented in the Appendix.

Finally, 48 filler trials were generated with the same distribution of target and context word frequencies; thus, of the 48 word trials in the experimental session, 12 (25%) used a related context (six forward and six backward). For subjects in the legal nonword condition, 12 (25%) of the nonword trials were pseudoprimes (e.g., DOCTOR gurse).

Procedure

The subjects ran in two sessions on consecutive days. The first was a practice session and was used to familiarize the subjects with the task and to find a target-exposure duration that produced roughly 25% errors. Related primes did not occur in the practice session, and none of the practice stimuli were repeated in the experimental session. Half of the 110 trials were words and half were nonwords. The first 2 trials were used to familiarize the subject with the experiment; the remaining 108 trials were divided into 6 blocks of

18, with exposure duration adjusted, if necessary, at the end of each block. The lexical decision task was explained, as were the characteristics of the stimulus presentation. The subjects were told that the pretarget prime was a cue to the location of the upcoming target and were asked to fixate it, although they were not informed that related words would appear. The subjects were also asked not to look up from the display until they had made their response, in order to avoid inadvertent experimenter cues.

Finally, response bias was explained and the subjects were asked to try to avoid bias. If the subjects made twice as many errors on nonword trials as on word trials, or vice versa, over the course of 36 trials (two blocks), they were informed that they were showing a "word" bias or a "nonword" bias. They were then asked to be either stricter or more liberal, as appropriate, in making a word response when uncertain. This bias feedback was not provided during the experimental session.

The rare subject who was unable to achieve binocular fusion on the pretarget primes, or was unable to report the stimulus to either eye at long durations, was paid for one session and was not run in the experimental session.

The experimental session was divided into five blocks. The first block consisted of 24 practice trials (half words, half nonwords) and was used to make additional adjustments in target-exposure duration. Each of the four experimental blocks contained 18 trials, with an equal proportion of filler versus experimental trials and an equal number of trials at each level of the three within-subject factors (although not an equal number of trials from each cell of the design). Between blocks, the subjects were allowed a brief rest period and exposure duration was adjusted if the total error rate in the preceding block fell outside a 16%-33% error range. In Experiments 1 and 2 of Stone (1985), the legal and illegal nonword conditions were equated by fixing the target-mask onset asynchrony at 67 msec for all subjects in both conditions. Since neither means of equating the two conditions (equal performance levels vs. equal stimulus-onset asynchronies) is unequivocally preferable, it was important to demonstrate that the same results were obtained with both methods of equating the two nonword-lexicality conditions.

Each trial began with a 1.5-sec binocular presentation of the pretarget prime. Immediately upon prime offset, the target appeared at the same location in the left eye for a duration that was set for each subject by the previously described method. Immediately upon offset of the target, the posttarget context appeared at the same location in the right eye for 1.5 sec. The display was dark between trials. After each trial, the subjects were told by the experimenter whether their response was correct or an error.

Eight presentation orders were crossed in a Latin square design with the between-subject control factors assigning stimulus items to experimental conditions. These presentation orders were obtained using two factors: permutations of block order (four levels) and reversals of presentation order (two levels). As a result, each item and each experimental condition had the same mean position in the presentation order across the full between-subjects design.

Design

There were four fully crossed experimental conditions, each with two levels, giving 16 conditions:

1. Nonword lexicality was a between-subjects factor, with half the subjects receiving only legal nonwords and half receiving only illegal nonwords. The same word targets and context words were used in the two conditions.

2. Prime direction was a within-subjects factor, with half of the experimental trials for each subject run in the forward-prime condition and half in the backward-prime condition.

3. Prime relatedness was a within-subjects factor, with half of the experimental trials for each subject run in the semantically related condition and half in the unrelated condition.

4. Target type was a within-subjects factor, with half of the trials for each subject run in the word-target condition and half in the nonword-target condition.

Each item appeared once per subject but was run equally often in each condition across subjects.

The five planned comparisons are described in the Results section. Omnibus ANOVAs for subjects and items were also performed to check for unexpected results.

RESULTS

Error scores for each experimental condition are presented in Table 1.

The primary analyses tested forward versus backward priming in the illegal nonword versus the legal nonword conditions. In particular, they tested for the interaction found in the first two experiments of Stone (1985), in which prime facilitation was equivalent in the forward-prime/legal nonword, the forward-prime/illegal nonword, and the backward-prime/illegal nonword conditions, but was significantly less in the backward-prime/legal nonword condition. This comparison was used because it provides a more stringent test of replication than the standard test for an interaction.

In a first test, priming was measured by taking the difference between errors in the related-prime and the unrelated-prime conditions. Since this measure of priming facilitation could be influenced by significant differences between the unrelated controls, a second test directly compared errors in the unrelated-prime condition.

The first test (using the difference in errors for related and unrelated primes) was significant [$F(1,46) = 11.074$, $MS_e = 0.0146$, $p < .05$, for subjects; $F(1,47) = 6.283$, $MS_e = 0.0516$, $p < .05$, for items]. The backward-prime/legal nonword condition produced significantly less

Table 1
Mean Error Scores by Condition

	Words		Nonwords	
	Forward Prime	Backward Prime	Forward Prime	Backward Prime
Legal Nonword Condition				
Related	0.188	0.285	0.229	0.174
Unrelated	0.313	0.299	0.174	0.188
Facilitation	0.125	0.014	-0.055	
Illegal Nonword Condition				
Related	0.201	0.208	0.236	0.278
Unrelated	0.313	0.299	0.236	0.271
Facilitation	0.112	0.091		

priming than did the other three conditions. The second test (for a significant difference in baseline performance in the legal and illegal nonword conditions) was not significant [$F(1,46) < 1.0$, $MS_e = 0.0092$, for subjects; $F(1,47) < 1.0$, $MS_e = 0.0467$, for items].

A third test checked for a significant difference in response bias in the legal nonword and illegal nonword conditions, using the interaction of nonword lexicality and target type for unrelated contexts. This comparison was included because a difference in response bias could be important for spreading-activation theories (see Discussion). This test was not significant [$F(1,46) = 2.787$, $MS_e = 0.0114$, $p > .10$, for subjects; $F(1,47) = 2.040$, $MS_e = 0.0313$, $p > .10$, for items].

A final pair of tests checked the robustness of the priming-bias effect reported by Schvaneveldt and McDonald (1981). They found that in masked lexical decisions, the benefit due to a related context for word trials (e.g., DOCTOR/nurse) was negated by an equivalent cost due to a related context on nonword trials (e.g., DOCTOR/nerse). Replication of the priming-bias effect was of interest because it played an important part in the development of the resonance model of Grossberg and Stone (1986). As noted previously, the pseudoprime targets in this experiment were constructed to be as discriminable as possible from their source words, within the constraint of changing a single letter. Thus, a test for an interaction of context type and target type for the forward-prime/legal nonword-foil condition would indicate if the priming-bias effect occurs under less favorable conditions. A test for a main effect of context type (related vs. unrelated) for the forward-prime/legal nonword-foil condition was also included to test for a significant gain in sensitivity due to priming.

The bias test was significant [$F(1,23) = 8.543$, $MS_e = 0.0229$, $p < .01$, for subjects; $F(1,47) = 7.201$, $MS_e = 0.0543$, $p < .05$, for items]. The 12.5% benefit for related-word trials was compromised by the 5.5% cost for related-nonword trials.

The sensitivity test was not significant [$F(1,23) = 1.200$, $MS_e = 0.0241$, $p > .10$, for subjects; $F(1,47) = 1.290$, $MS_e = 0.0579$, $p > .10$, for items]. These results support the contention that under pattern-masking conditions, semantic priming affects primarily response bias rather than sensitivity to lexical status, although the trend toward less than complete bias suggests that the priming-bias effect is sensitive to the similarity of pseudoprime nonwords to their source word (e.g., DOCTOR/nerse vs. DOCTOR/gurse). A discussion of the priming-bias effect in the context of expectancy-set versus spreading-activation theories can be found in Schvaneveldt and McDonald (1981). A discussion in terms of resonance models can be found in Grossberg and Stone (1986).

Excluding sources that compared words with nonwords, the omnibus ANOVA produced only a main effect of prime relatedness, with fewer errors on related-prime trials [$F(1,46) = 6.236$, $MS_e = 0.0205$, $p < .05$, for subjects; $F(1,47) = 6.370$, $MS_e = 0.0401$, $p < .05$, for items].

Finally, the pattern of priming effects was checked using LSD tests generated from the error terms of the omnibus ANOVAs using subjects and items as the random factor. Related primes were compared with unrelated primes for each prime direction \times nonword type condition. Both forward-prime conditions and the backward-prime/illegal nonword condition showed significant prime facilitation ($ps < .05$ on all six tests). The difference between related and unrelated primes for the backward-prime/legal nonword condition was not significant. Furthermore, performance for related primes was not significantly different for forward and backward primes in the illegal nonword condition ($ps > .10$ for both subjects and items), but was significantly different for forward and backward primes in the legal nonword condition, with fewer errors in the forward-prime condition ($ps < .05$ for both subjects and items).

DISCUSSION

Backward priming equivalent to forward priming can occur in the bidirectional priming paradigm. However, significant backward priming occurred only when the nonword foils were illegal. This is the same pattern of results found in the first two experiments of Stone (1985), and does not support any of the "most natural" predictions of spreading-activation and expectancy-set theories.

It is worth noting that in a similar paradigm, Briand, den Heyer, and Dannenbring (1988) also looked at backward priming of masked lexical decisions and found significant backward priming with legal nonword foils and prime-target onset asynchronies up to 1 sec. Although their results appear to be in conflict with those of the present study (including their failure to replicate the priming-bias effect of Schvaneveldt & McDonald, 1981), an important difference in methodology could be responsible for this difference.

In the Briand et al. (1988) study, subjects saw a prime, which was masked by a row of asterisks. This prime was either a word or a nonword. They were then presented a target word at varying prime-target onset asynchronies. The subjects were told that their primary task was to name the target. A subsequent, secondary lexical decision was then made concerning the prime. Briand et al. found significant backward priming of this lexical decision, which did not interact with prime-mask onset asynchrony or with prime-target onset asynchrony.

Under those conditions, it is quite likely that information about the prime was held in postlexical short-term storage while the subjects focused on the primary task of naming the target. If this interpretation is correct, the backward priming that Briand et al. (1988) observed is a different phenomenon from the backward priming due to a mask word that is related to the lexical decision target when the lexical decision is the focal task. However, the possibility that lexical processing can enhance information held in postlexical short-term storage is intriguing, particularly considering that backward priming in

their paradigm was apparently not due to increased bias toward word responses.

We consider first the implications of our results for the assumption that the nonword-lexicality manipulation shifts the emphasis given to lexical and nonlexical sources of information in the bidirectional priming paradigm. We then consider implications for spreading-activation and expectancy-set theories and, finally, we suggest an alternative to node theories (in the broad sense) which provides a natural account of the results by combining the important insights of shared activation (from spreading-activation theories) and matching of expectations against inputs (from expectancy-set theories).

Lexical and Nonlexical Sources of Information

Shulman and Davidson (1977) argued for greater reliance on nonlexical processes (e.g., a pronunciation subsystem) when illegal nonwords are used as foils, on the basis of reduced priming effects with illegal nonword foils.

In the present study, the magnitude of forward priming was relatively unaffected by the nonword-lexicality manipulation, whereas the pattern of backward priming was the reverse of what would be expected given Shulman and Davison's (1977) analysis (there was more priming in the illegal nonword condition). This suggests that lexical processing was employed equally in both nonword-lexicality conditions.

Spreading-Activation Theories

Backward priming effectively equivalent to forward priming (illegal nonword condition) requires a very rapid spread of activation between word nodes (as in Wickelgren, 1976). This limit on the rate of spread could be raised somewhat if the nonsignificant (2.1%) difference between forward and backward priming in the illegal nonword condition is, in fact, reliable.

However, this possibility assumes little or no controlled priming due to the pretarget prime. Inhibition from the unrelated pretarget prime in the backward-prime condition could partially counteract facilitation due to the semantically related posttarget prime. In nonmasked lexical decisions with forward priming, reduction of the proportion of related contexts does not eliminate controlled priming. Rather, the degree of controlled priming decreases as the proportion of related contexts decreases (Tweedy et al., 1977). Thus, a slight disadvantage for backward priming, relative to forward priming, might be due to controlled inhibition by the pretarget prime in the backward-priming condition.

The loss of backward priming in the legal nonword condition poses a more difficult problem for spreading-activation theories. The problem arises from the observation that the stimulus event of interest (word trials) in the two nonword-lexicality conditions is identical. This implies that performance on a given trial cannot depend solely upon analysis of the current stimulus. Performance

must also be affected by what *might* have occurred as the target.

An obvious method of affecting performance in such a manner is to adjust the threshold for word nodes as a function of nonword lexicality. When a node's threshold is lowered, it is more likely to exceed threshold before facilitation from the backward prime can influence processing. This would, in turn, reduce the magnitude of backward priming.

However, such an adjustment has another important consequence for performance. All else being equal, a decrease in thresholds will increase bias toward word responses, since more trials—both word and nonword—will produce a suprathreshold word node, and thus a word response. Assuming that legal nonwords produce greater lexical activity than do illegal nonwords, a fixed threshold will produce greater false alarms to legal nonwords than to illegal nonwords. Given no observed difference in response bias for the two nonword conditions (as noted in the Results section), thresholds would need to be *higher* in the legal nonword condition than in the illegal nonword condition.

The adjustment of thresholds needed to account for the observed pattern of response bias and the adjustment of thresholds needed to account for the pattern of backward-prime facilitation are in conflict. If thresholds are lower in the illegal nonword condition, it is possible to obtain equivalent response bias in the two nonword conditions, but backward-prime facilitation would then be more effective in the legal nonword condition. If thresholds are lower in the legal nonword condition, backward-prime facilitation would be less effective in the legal nonword condition, but there would also be greater bias toward word responses in the legal nonword condition.

Note that, given the alternative assumptions concerning spreading activation that have been entertained, only the observed result—backward priming given illegal nonword foils, but not given legal nonword foils—is incompatible with spreading-activation theories.

Expectancy-Set Theories

The results of this experiment violate the joint assumptions that masking prevents verification, and that verification is the sole locus of context effects (Paap et al., 1982). These assumptions, taken together, suggest that neither forward nor backward priming should be facilitative in the bidirectional priming paradigm. Even if accuracy on word trials reflects the likelihood of a match before the mask terminates verification, so that forward priming could result from an earlier search of the semantic search set, backward priming should not occur because verification of the target is terminated as soon as processing of the mask begins.

Because the assumption that masking terminates verification is not fundamental to expectancy-set theories, we considered an alternative assumption concerning the effect of masking—that masking disrupts, but does not ter-

minate, verification. In this case, verification would match lexical schemata from its search set with a degraded, decaying perceptual schema for the target. Schemata appearing early in the search sequence would be matched against more complete perceptual schema than would those appearing late in the sequence, and would thus suffer fewer errors.

Backward priming would occur if a semantic search set could be generated for the posttarget prime and be employed before search using the visual search set. This would presumably entail verification of the posttarget prime, as well as generation of its semantic search set, before verification of the target. This is particularly true since the effective equivalence of forward and backward priming implies that the semantic search sets for pre- and posttarget primes are searched equally early, on average. Given a rate of decay for the degraded perceptual schema sufficient to produce the observed magnitude of facilitation, a delay of target verification until a search set had been generated for the posttarget prime would entail a significant loss of target information. Such a processing strategy would presumably introduce costs that might even outweigh the benefits in the backward-priming condition (a relatively uncommon condition in the experiment). However, processing strategies are not always optimal (Stroop, 1935). It is quite possible that recognition of a meaningful mask (posttarget prime) is obligatory, despite the costs.

The primary problem with this argument is the loss of backward priming when the foils are legal nonwords. Of course, one could argue that the search strategy in the verification stage was different in the legal nonword condition (i.e., a semantic set generated by the backward prime was less likely to be available early in verification). A principled account of this difference in strategy would be that the longer stimulus-onset asynchronies in the legal nonword condition (needed to equate performance in the two conditions) made delay of target recognition until a search set had been generated for the posttarget prime prohibitively detrimental. However, this explanation is unlikely, given that the same pattern of results was obtained when onset asynchronies were equated in the two conditions (Stone, 1985). Of course, it is possible that nonword lexicality affects the generation of a semantic search set for some other reason; however, it is not clear to us why this would be so, except to account for our results.

An Alternative to Node Theories

These results raise problems for both spreading-activation and expectancy-set theories, which share the assumption of structurally unitized word representations (or nodes, in the broad sense). We suggest that an alternative approach, in which word representations are *functionally unitized*, and can combine the important insights of shared activity between related representations (from spreading-activation theories) and of active matching of expectations against input (from expectancy-set theories).

This approach can provide access to a lexicon organized in terms of semantic space.

We begin by noting that verification can be defined broadly as a two-phase recognition process with a passive, stimulus-driven initial phase and an active second phase, in which expectations derived from knowledge in long-term memory are matched against the stimulus-generated representation (Broadbent & Broadbent, 1980). A number of studies have provided supporting evidence for this general framework (Becker, 1979, 1980; Becker & Killion, 1977; Broadbent, 1977; Broadbent & Broadbent, 1980; Schvaneveldt & McDonald, 1981; Van Orden, 1987, in press). This framework does not require serial search. Indeed, adaptive resonance (Carpenter & Grossberg, 1987; Grossberg, 1980; Grossberg & Stone, 1986; Stone, 1986) is a form of verification that performs matching in parallel.¹ Before developing this resonance framework, we shall briefly discuss the semantic-space approach to semantic memory that it assumes.

Semantic-Space Theories

Consider a high-dimensional space in which each dimension represents some semantic feature.² Each word (or concept) is represented by a pattern of activations (or strengths) for each feature. Related words *share* activation on a suite of features. Thus, network theories and semantic-space theories both assume that representations for related words share activation. They differ in that network theories assume spatially distinct concept nodes, so that shared activation must spread via links between these nodes. In semantic-space theories, activation is shared at the same feature unit and thus need not spread to become manifest. However, a word's representation is now distributed across units—no single unit's activation expresses the likelihood that a particular word is currently being represented.

Semantic-space theories have been popular in the study of semantic processing (Rosch, 1978; Smith, Shoben, & Rips, 1974); however, they have been conspicuously absent in the study of lexical access. This absence has been due, we believe, to the serious problem of selecting a single, unitized representation for a lexical stimulus when representations are *patterns* of activation (or feature strengths) across units (or dimensions).

In postlexical processing, one can assume that a single-word representation has been selected. However, lexical access must deal with the strongly correlated visual structure of words. Presentation of one word (e.g., MAN) will provide some evidence for visually similar words (e.g., PAN and CAN). Consequently, the lexical representations for many words will be activated by the presentation of a single word. The problem thus emerges of selecting the correct lexical representation from among many candidate representations. Structural unitization (nodes in the broad sense) makes selection of a single such representation easy since a word's representation is "all in one place."

In spreading-activation theories, a single node crosses a threshold and is thus selected. However, thresholds will not work with distributed representations—one cannot simply select the most active features for two reasons. First, less active (or weaker) semantic features may be important to a representation (Smith et al., 1974). Second, relatively strong features from other representations might not be excluded (in the above example, MAN partially activates PAN and CAN, which could jointly produce strong activation of the feature “metallic”). Some form of “clean-up” process is required to eliminate unwanted representations in the encoding.³

A solution to this dilemma lies in the use of a verification process to unitize word representations *functionally*. Functional unitization occurs when components of a representation act together as a coherent whole because of processing dynamics. Each component helps support the others, so that the components increase or decrease in activation together. In this way, the collection of components *acts as if* it were a single unit.

Resonant Verification

Our discussion of resonant verification is organized around several essential properties. We have chosen this approach because discussion in terms of such *design principles* helps to clarify why and how a formal model operates as it does. Design principles are constraints on system design derived from empirical observations within the constraints of a chosen framework. They serve as intermediaries between empirical observation and the development of formal models (Stone, 1988; Van Orden, Pennington, & Stone, 1989). Many models may satisfy the same design principle. This design principle clarifies which aspects of the models allows them all to account for the empirical observations upon which the design principle is based.

In this approach to theoretical methodology, the addition of new design principles narrows the set of admissible models. For example, verification is a design principle that can be realized in models using either schemata or distributed patterns of activity as representations. The addition of functional unitization as a design principle narrows the set of admissible models to those with distributed patterns as representations.

The assumption of distributed representations during word recognition necessitates a distinction between overall patterns of activation and the pattern of activation that corresponds to a single, coherent interpretation. We emphasize this by distinguishing between encodings and representations. An encoding is the overall pattern of activity resulting from a stimulus at a given time. A representation is a pattern of activation that represents a single, coherent interpretation of a stimulus event. Thus, the encoding of a stimulus during lexical selection may consist of the representations for a number of candidate words. Lexical selection is the process by which an encoding comes to consist of a single representation.

Consider now two levels of processing: a visual level, at which a pattern of activity constitutes an orthographic encoding in terms of spelling features, and a lexical level, at which a pattern of activity constitutes a lexical encoding, for our purposes, in terms of semantic features.⁴ The visual similarity of two words is reflected in shared subpatterns at the visual level. The semantic similarity of two words is reflected in shared subpatterns at the lexical level.

Presentation of a lexical stimulus (e.g., a word) produces a pattern of activity at the first, visual level, which encodes its spelling features. As in the logogen model (Morton, 1969), connections run from the “feature detectors” at the visual level to units at the lexical level, and activation is passed along these connections, exciting lexical representations. However, these lexical representations are distributed patterns of activity rather than discrete, “word-detector” nodes. The total strength of activation for the lexical representation of a word is proportional to that word’s visual similarity to the stimulus, just as is the activation of a word node in the logogen model. The logogen’s activation has simply been distributed among the components of a lexical feature pattern.

As in expectancy-set theories, an expected visual representation for each lexical representation is matched against the stimulus-driven visual encoding. However, this resonant verification differs from serial-search verification in several important ways.

First, matching of top-down expectations against bottom-up input occurs in parallel for each active lexical representation, rather than serially for one lexical representation at a time. Furthermore, matching is performed at the visual level, at which top-down connections from the lexical level to the visual level literally add top-down expected patterns to the bottom-up stimulus pattern. Goodness of match depends upon activation from both sources: visual features receiving both top-down and bottom-up support are more likely to be strengthened by the matching process, whereas visual features receiving support from a single source are more likely to be weakened by the matching process (Carpenter & Grossberg, 1987). In other words, resonant matching is a constructive process, in that the visual encoding is altered during matching on the basis of the agreement between bottom-up and top-down sources of information.⁵

Second, multiple matching cycles are performed. In each cycle, lexical representations whose expectations produce stronger matches at the visual level become more active (since matching has strengthened their bottom-up support from the visual level), whereas lexical representations whose expectations produce weaker matches become less active (since matching has weakened their bottom-up support). Likewise, components of the visual encoding are successively strengthened or weakened in each successive matching cycle on the basis of their participation in a match. Each cycle consists of a feed-forward passing of activation from the current visual encoding to the lexical level to produce a new lexical encoding, a feed-

backward passing of activation based on the expectations of the representations in this new lexical encoding to the visual level, and a matching of top-down and bottom-up patterns at the visual level to produce a new visual encoding. Over cycles, lexical representations that are inconsistent with the bottom-up visual encoding are suppressed until a single best interpretation wins. Thus, selection in resonant verification is based on the *best* match, rather than *first* match found, as in serial-search verification.

Third, unitization of the word's representation arises in the course of this verification process. The mutually supportive feedback loop between levels ties together the visual and lexical representations, as well as the component activations of each representation. This occurs because of a fundamental property of processing-distributed representations. Each component of a lexical representation contributes to the *whole* expected pattern used in matching at the visual level. This expected visual pattern (in strengthened or weakened form) then provides feed-forward support for the whole lexical representation. Thus, through the verification loop, each component of a lexical representation affects the fate of all other components in the representation: a component whose contribution to the top-down expectation increases the match at the visual level will increase the support at the visual level for the whole lexical representation of which it is a part. As a result, the correct unit of analysis is not a loose collection of semantic features; rather, it is the *resonance* that emerges from the verification process—the lexical and visual representations along with their mutually supportive feedback loops.

Fodor and Pylyshyn (1988) have argued that connectionism is fatally flawed because the joining of parts into a whole requires a higher order node connected to the nodes for the components. Yet resonant verification (a "connectionist" approach) *does* join parts into a whole without the use of a higher order node. The componentiality problem noted by Fodor and Pylyshyn is a serious problem. However, it is not a problem for "connectionist" models in general but, rather, for models (symbolic or connectionist) that employ simple associations between structurally unitized representations.

This more complex view of verification arose from questions about how unitization arises during the course of learning. Serial-search mechanisms have been successful in accounting for the recognition of words by skilled readers under relatively "normal" circumstances because, under these conditions, the above distinctions are often unimportant for a qualitatively correct description of performance (e.g., the best match will also be the only, and thus the first, match found in a serial search). However, when issues such as degree of unitization (Drewnowski & Healy, 1977; Healy, Conboy, & Drewnowski, 1989) are important, or when unusual conditions exist (as in bidirectional priming), the above distinctions can be critical.

Bidirectional Priming and Resonant Verification

Our account of bidirectional priming begins with a general description of priming within the resonant-verification framework, then considers the effects of nonword lexicality on processing strategy. This discussion will provide the additional principles needed to understand why backward priming should be sensitive to the choice of nonword foils.

Semantic priming cannot be understood solely in terms of shared semantic features. Somehow, the increased activation of shared semantic features due to priming must benefit the *whole* target representation. Otherwise, we simply have the sum of two patterns that happen to overlap. It is functional unitization that allows facilitation of a subpattern to facilitate the whole target representation.

Consider the initial matching cycle. The representation of a prime will still be present at the lexical level. The portion of a target word's lexical representation that overlaps with that of a related prime will be more active than it would have been without the related prime. Since each component of the lexical representation contributes to the *whole* top-down expectation, the target's correct top-down expectation is slightly strengthened. This improves the relative match between the target's correct top-down expectation and the stimulus-driven visual encoding. In the next cycle, the target's *whole* lexical representation will thus receive greater bottom-up support. Note, however, that the prime's lexical representation is weakened, since its expectation poorly matches the visual encoding. Facilitation of target recognition results from this boost to its lexical representation during early verification cycles. On later cycles, the prime's lexical representation will have been suppressed due to mismatch with the visual encoding.

Nonword-lexicality effects reflect differences between the lexical encodings for illegal versus legal nonwords. Illegal nonwords are not highly similar to many (if any) words. Consequently, their initial lexical coding produces little overall lexical activity spread across many word representations with very low levels of activation. Because candidate lexical representations are weak and produce poor matches with the visual encoding, no resonance forms and the lexical encoding rapidly collapses (a system that always selects the best match is clearly unacceptable: the best match must effectively exceed some match criterion to induce resonance). Legal nonwords, on the other hand, produce significant activation of lexical representations for similar words. More matching cycles are required to suppress the lexical representations of words that only moderately mismatch the nonword's visual encoding.

These differences suggest a difference in processing strategy based on a characteristic of resonance matching not found in all-or-none serial matching—the relative weight given to top-down and bottom-up information during matching.

Increasing emphasis (weight) on bottom-up information means that top-down expectations will have less effect on

the visual encoding that results from matching. Resonant verification then resembles a slow resolution of initial bottom-up support. Lexical representations whose expectations almost match the stimulus-driven visual encoding will eventually be suppressed because bottom-up evidence of the small mismatch will persevere. However, a lexical representation close enough to accept a match will be pulled into resonance slowly, because it has little ability to draw the visual encoding toward its expectation. The result is a higher *effective* match criterion⁶ when the weight given to bottom-up information is increased.

Increasing emphasis (weight) on top-down information means that close matches can strongly influence the visual encoding after matching. This alteration of the visual encoding both amplifies the increase in strongly matching representations and hastens the decline of weakly matching representations. As a result, poor matches are suppressed more quickly. However, mismatches that are close to matches are more likely to initiate a false resonance, leading to false positive recognition. Their strong initial activations generate strong top-down expectations, which amplify their supporting bottom-up evidence. The result is a lower effective match criterion when the weight given to top-down information is increased.

In sum, the effective match criterion can be shifted by “tuning” the relative emphasis on top-down and bottom-up information in matching: increasing the emphasis on bottom-up evidence raises the effective match criterion, whereas increasing the emphasis on top-down expectations lowers the effective match criterion. However, unlike adjustments of a threshold-style match criterion, adjustments of this effective match criterion affect processing at all degrees of match/mismatch. The stricter the match criterion, the slower the suppression of noise and weakly matching expectations, and thus the slower the formation of a resonance. This “tuning” of the matching process allows the system to adjust for a variety of task demands (e.g., strong bottom-up in proofreading, in which close orthographic mismatches must be detected, vs. strong top-down in skimming, in which close orthographic mismatches are a distraction).

In the bidirectional priming paradigm, the use of a backward-pattern mask places a premium on the tradeoff between the speed of resonance formation and the strictness of the match criterion. In any given condition, the balance between top-down and bottom-up information in the matching process can have a major impact on performance. The greatest possible difference in the lexical activity induced by words and nonwords must be achieved by the time masking disrupts the recognition process. We consider first the optimal settings of this balance when nonwords are illegal or legal, and then consider the implications of these settings for forward and backward priming.

In all cases, the limited time before mask onset demands rapid suppression of noise and weak matches for acceptable performance. Masking conditions thus suggest a low match criterion through emphasis on top-down informa-

tion. The cost of this tuning is a high rate of false positive recognition. However, the system can minimize this cost by changing the relative emphasis on top-down and bottom-up information during the course of recognition.

In the early matching cycles, lexical representations are not yet strongly activated, since they grow because of increasing support from matches on many cycles. Emphasis on top-down information in these early matching cycles allows suppression of noise and relatively weak matches, without major alteration of the visual encoding toward the closer matches, since expectations are still weak in absolute terms. However, the lexical representations whose expectations closely match the visual encoding are becoming more strongly activated. On later cycles, this would lead to excessive alteration of the visual encoding toward the closest matches. The system can avoid this problem by raising the match criterion (through increasing emphasis on bottom-up information) as the number of matching cycles increases. Because of the accelerated suppression of noise during the early matching cycles, the cost of raising the match criterion for later cycles—slower suppression of spurious lexical codes—is reduced. In this later phase of processing, the degree of emphasis on bottom-up information (and thus the match criterion) can be set to optimize discrimination between correct matches and false positive matches.

The setting of the match criterion in this later phase depends upon the type of nonword foils. Because illegal nonwords generate relatively weak initial encoding spread among many lexical representations, verification can continue into later matching cycles with a low match criterion. A slow shift away from a top-down emphasis maintains a rapid suppression of poor matches. As a result, the lexical encoding of an illegal nonword collapses more rapidly than it would if the emphasis shifted more rapidly toward a bottom-up emphasis. Furthermore, the lexical representation of a word stimulus will achieve a stronger resonance (and activation) by mask onset with a slow shift toward bottom-up emphasis, due to its increased ability to alter the visual encoding in its favor.

Legal nonwords, on the other hand, can generate lexical encodings similar to those generated by words, so that a rapid shift from a top-down emphasis to a bottom-up emphasis is required to avoid excessive false positive recognitions. However, word-target representations achieve weaker resonance (and activation) by mask onset since they are less effective at drawing the visual encoding toward their expectations.

Recall that prime facilitation resulted from increased activity for the lexical subpattern shared by the target and the prime. This shared activity contributes to activation of the whole target representation via its effect on the visual encoding that results from matching, and is thus sensitive to the relative weight given to top-down versus bottom-up information in matching. When more emphasis is given to top-down expectations, they have a stronger effect on the visual encoding that results from matching, so that the expectation due to the shared subpattern is more

strongly reflected in the visual encoding after matching. Thus, its indirect support of the whole lexical representation is enhanced. On the other hand, more emphasis on bottom-up information means little effect of the subpattern's expectation of the visual encoding (especially since, as a partial representation, it provides only a partial match). As a result, a shared subpattern will have little effect when matching emphasizes bottom-up information.

Forward priming is effective in both nonword conditions because it has its effect on early matching cycles, when top-down information is emphasized in matching. Backward priming occurs in a later phase of processing, when the weighting of top-down and bottom-up information differs in the two nonword conditions. In the illegal nonword condition, the shared subpattern affects processing when there is still an emphasis on top-down information in matching. Because facilitation occurs over only a few matching cycles, forward and backward priming will have roughly equivalent effects. In the legal nonword condition, the shared subpattern affects processing when an emphasis on bottom-up information makes it fairly ineffective in contributing to the activation of the target representation. Indeed, it is not surprising that, during a phase of processing when stimulus similarity is a hazard to correct judgments, the similarity of target and prime lexical representations produces little or no facilitation.

CONCLUSION

Spreading-activation theories incorporate the important idea of shared activation but fail to account for both the equivalence of forward and backward priming given illegal nonwords as foils and the loss of backward priming given legal nonwords as foils. Expectancy-set theories incorporate the important idea of matching top-down expectations against visual encodings but can only accommodate the above results with ad hoc assumptions. Resonant verification brings together the important ideas from each approach.

The properties of resonant verification derive from consideration of a wide range of empirical phenomena, since the ultimate value of resonant verification will depend upon its successful application to a wide range of phenomena in word recognition. However, in the above discussion, we employed only those properties of resonant verification necessary for understanding the results from the bidirectional priming paradigm. Several important issues were not discussed, such as the role of phonology (Van Orden, 1987, in press), the importance of the list-item error tradeoff (Grossberg & Stone, 1986; Stone, 1986) in comparing unmasked and masked recognition (especially as studies of backward priming in unmasked recognition indicate interesting differences from backward priming in the bidirectional priming paradigm, Kiger & Glass, 1983; Seidenberg, Waters, Sanders, & Langer, 1984), and the notion of subresonances (i.e., functionally unitized components of a larger resonance unit).

Furthermore, this framework for word recognition is still incomplete in its details. Several alternatives exist for formalizing some of these mechanisms. Further experimentation and analysis are necessary to choose between them. However, these future refinements must satisfy a methodological constraint that has, thus far, guided the development of this framework. That is, each design principle, and the mechanism that embodies it, must contribute to an understanding of a range of empirical phenomena—the use of highly specialized, single-purpose mechanisms is considered to be a sign of an undetected error in the development of the framework. If, for example, the shift of emphasis from top-down to bottom-up information during the course of recognition fails to have explanatory value for other experimental paradigms, its value in accounting for the results from the bidirectional priming paradigm will come to naught.

Regardless of the viability of this approach to theoretical development, the results of the bidirectional priming experiment reported here provide converging evidence against a spread of activation between word nodes in a passive-recognition system. Moreover, if this theoretical approach is accepted, these results also provide converging evidence against expectancy-set theories with serial-search verification. Granting the latter, these results can be taken as converging evidence that words are not represented by nodes (in the broad sense of structurally unitized representations) but, rather, that lexical representations are functionally unitized.

REFERENCES

- BATTIG, W. F., & MONTAGUE, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, **80**(3, Pt. 2).
- BECKER, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 556-566.
- BECKER, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 252-259.
- BECKER, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition*, **8**, 493-512.
- BECKER, C. A., & KILLION, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 389-401.
- BRIAND, K., DEN HEYER, K., & DANNENBRING, G. L. (1988). Retroactive semantic priming in a lexical decision task. *Quarterly Journal of Experimental Psychology*, **40A**, 341-359.
- BROADBENT, D. E. (1977). The hidden pre-attentive processes. *American Psychologist*, **32**, 109-118.
- BROADBENT, D. E., & BROADBENT, M. H. P. (1980). Priming and the passive/active model of word recognition. In R. Nickerson (Ed.), *Attention and performance* (Vol. 8, pp. 419-433). Hillsdale, NJ: Erlbaum.
- CARPENTER, G., & GROSSBERG, S. (1987). Neural dynamics of category learning and recognition: Attention, memory consolidation, and amnesia. In J. Davis, R. Newburgh, & E. Wegman (Eds.), *Brain structure, learning, and memory*. AAAS Symposium Series.
- CARROLL, M., & KIRSNER, K. (1982). Context and repetition effects

- in lexical decision and recognition memory. *Journal of Verbal Learning & Verbal Behavior*, **21**, 55-69.
- COLLINS, A. M., & LOFTUS, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, **82**, 407-428.
- COLLINS, A. M., & QUILLIAN, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior*, **8**, 240-247.
- DREWNOWSKI, A., & HEALY, A. F. (1977). Detection errors on the and: Evidence for reading units larger than the word. *Memory & Cognition*, **5**, 636-647.
- FISCHLER, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory & Cognition*, **5**, 335-339.
- FODOR, J. A., & PYLYSHYN, Z. W. (1988). Connectionism and cognition: A critical analysis. *Cognition*, **28**, 3-71.
- FORSTER, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. C. T. Walker (Eds.), *New approaches to language mechanisms* (pp. 257-287). Amsterdam: North Holland.
- GROSSBERG, S. (1980). How does the brain build a cognitive code? *Psychological Review*, **87**, 1-51.
- GROSSBERG, S., & STONE, G. O. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review*, **93**, 46-74.
- HEALY, A. F., CONBOY, G. L., & DREWNOWSKI, A. (1989). *Characterizing the processing units of reading: Effects of intra- and inter-word spaces in a letter detection task*. Unpublished manuscript.
- JACOBSON, J. Z. (1973). Effects of association upon masking and reading latency. *Canadian Journal of Psychology*, **27**, 58-69.
- JACOBSON, J. Z. (1974). Interaction of similarity to words, visual masks and targets. *Journal of Experimental Psychology*, **102**, 431-434.
- JACOBSON, J. Z., & RHINELANDER, G. (1978). Geometric and semantic similarity in visual masking. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 224-231.
- JAMES C. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 130-136.
- KIGER, J. I., & GLASS, A. L. (1983). The facilitation of lexical decisions by a prime occurring after the target. *Memory & Cognition*, **11**, 356-365.
- KUCERA, H., & FRANCIS, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- MCCLELLAND, J. L. (1978). Perception and masking of wholes and parts. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 210-223.
- MCCLOSKEY, M., & SANTEE, J. (1981). Are semantic memory and episodic memory distinct systems? *Journal of Experimental Psychology: Human Learning & Memory*, **7**, 66-71.
- MCKOON, G., & RATCLIFF, R. (1980). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning & Verbal Behavior*, **19**, 369-386.
- MORRIS, W. (Ed.) (1969). *The American heritage dictionary of the English language*. Boston: Houghton Mifflin.
- MORTON, J. (1969). Interaction of information in word recognition. *Psychological Review*, **76**, 165-178.
- NEELY, J. H., & DURGUNOGLU, A. Y. (1985). Dissociative episodic and semantic priming effects in episodic recognition and lexical decision tasks. *Journal of Memory & Language*, **24**, 466-489.
- PAAP, K. R., NEWSOME, S. L., McDONALD, J. E., & SCHVANEVELDT, R. W. (1982). An activation-verification model for letter and word recognition: The word superiority effect. *Psychological Review*, **89**, 573-594.
- PALERMO, D. S., & JENKINS, J. J. (1964). *Word association norms: Grade school through college*. Minneapolis: University of Minnesota Press.
- RATCLIFF, R., & MCKOON, G. (1981). Does activation really spread? *Psychological Review*, **88**, 454-462.
- ROSCH, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Categorization and cognition* (pp. 27-48). Hillsdale, NJ: Erlbaum.
- RUMELHART, D. E., HINTON, G. E., & WILLIAMS, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP group, *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1, pp. 318-362). Cambridge, MA: MIT Press.
- SCHVANEVELDT, R. W., & McDONALD, J. E. (1981). Semantic context and the encoding of words: Evidence for two modes of stimulus analysis. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 673-687.
- SEIDENBERG, M. S., WATERS, G. S., SANDERS, M., & LANGER, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition*, **12**, 315-328.
- SHOEN, E. J., WESCOURT, K. T., & SMITH, E. E. (1978). Sentence verification, sentence recognition, and the semantic-episodic distinction. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 304-317.
- SHULMAN, H. G., & DAVISON, T. C. B. (1977). Control properties of semantic coding in a lexical decision task. *Journal of Verbal Learning & Verbal Behavior*, **16**, 91-98.
- SHULMAN, H. G., HORNAK, R., & SANDERS, E. (1978). The effects of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory & Cognition*, **6**, 115-123.
- SMITH, E. E., HAVILAND, S. E., REDER, L. M., BROWNELL, H., & ADAMS, N. (1976). When preparation fails: Disruptive effects of prior information on perceptual recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 151-161.
- SMITH, E. E., SHOEN, E. J., & RIPS, L. J. (1974). Structure and process in semantic memory: A featural model of semantic decisions. *Psychological Review*, **81**, 214-241.
- SMOLENSKY, P. (1988). On the proper treatment of connectionism. *Behavioral & Brain Sciences*, **11**, 1-74.
- SPELTING, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74**, 1-29.
- STONE, G. O. (1985). *Semantic facilitation by a word mask: Implications for theories of word recognition and interruption masking*. Unpublished doctoral dissertation, University of California, San Diego.
- STONE, G. O. (1986, August). *A distributed model of semantic priming*. Paper presented at the 19th Annual Meeting of the Mathematical Psychology Society, Boston, MA.
- STONE, G. O. (1988). From data to dynamics: The use of multiple levels of analysis. *Behavioral & Brain Sciences*, **11**, 54-55.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- TAYLOR, G. A., & CHABOT, R. I. (1978). Differential backward masking of words and letters by masks of varying orthographic structure. *Memory & Cognition*, **6**, 629-635.
- TURVEY, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, **80**, 1-52.
- TWEEDY, J. R., LAPINSKI, R. H., & SCHVANEVELDT, R. W. (1977). Semantic-context effects on word recognition: Influence of varying the proportion of items presented in an appropriate context. *Memory & Cognition*, **5**, 84-89.
- VAN ORDEN, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, **15**, 181-198.
- VAN ORDEN, G. C. (in press). Phonologic mediation is fundamental to reading. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*. Hillsdale, NJ: Erlbaum.
- VAN ORDEN, G. C., PENNINGTON, B., & STONE, G. O. (1989). *Word identification in reading and the promise of subsymbolic psycholinguistics*. Manuscript submitted for publication.
- WICKELGREN, W. (1976). Network strength theory of storage and retrieval dynamics. *Psychological Review*, **83**, 466-478.

NOTES

1. Adaptive resonance can also perform serial matching (Carpenter & Grossberg, 1987). However, the serial-matching mode occurs primar-

ily during learning. Because words are very well learned, we consider only its parallel-matching mode.

2. These semantic features need not correspond to our a priori notions of the most meaningful components of a concept. In fact, one can obtain the greatest power from distributed processing if the units are "microfeatures," or components at a smaller "grain size" than the smallest meaningful components (Smolensky, 1988; Van Orden, Pennington, & Stone, 1989). In any case, our analysis requires only that shared features correspond to shared meaning.

3. The most common solution to this problem of "cross-talk" in connectionist theories is to eliminate the unwanted representations through supervised learning, so that they do not occur for well-learned stimuli (Rumelhart, Hinton, & Williams, 1986). However, if this were the only means of cleaning up encodings, a number of important factors (e.g., familiarity) would have no effect for well-learned stimuli such as words.

If elimination of cross-talk through learning is incomplete for very familiar stimuli, such as words, there is still a need for some kind of clean-up process during recognition.

4. This simple configuration of levels is the minimum needed for our analysis. For discussions of more complex configurations in this framework, see Grossberg and Stone (1986) and Van Orden (1987).

5. Resonant verification does more than identify the stimulus; it is also the process responsible for suppressing noise and regularizing variable inputs (e.g., different type fonts). Indeed, it originated as a mechanism for generalization and categorization.

6. Resonant verification does not use an explicit match criterion. Suppression of mismatch emerges from the interaction of the component processes of resonant verification. Thus, we refer to the strength level below which suppression occurs as the *effective* match criterion.

APPENDIX EXPERIMENTAL STIMULI

Yoked Pair	Targets			Primes	
	Word	Illegal	Legal	Related	Unrelated
1	week	ekwe	weel	days	room
	five	efvi	mive	four	case
2	hand	ndah	hane	feet	line
	knob	obkn	knop	door	name
3	sail	iasl	saip	boat	rule
	beer	ebre	meer	wine	hill
4	beef	ebfe	beel	meat	asia
	wash	hasw	wast	dirt	pace
5	farm	fmra	farl	barn	belt
	moon	mnoo	mool	star	pack
6	bird	rdib	bire	hawk	mist
	tree	etr	dree	pine	yarn
7	most	tms0	mose	least	young
	dark	rdak	sark	light	thing
8	baby	ybba	saby	child	start
	song	gsno	fong	music	level
9	lies	lsei	nies	truth	press
	leaf	afle	neaf	plant	blood
10	rock	rkco	gock	stone	smile
	book	okbo	bood	novel	india
11	lion	ionl	liod	tiger	relic
	fork	kfor	gork	spoon	sword
12	game	mgea	gamp	chess	value
	rose	eors	bose	thorn	yeast
13	south	ohstu	pouth	north	class
	white	wtehi	chite	black	peace
14	dance	cnade	rance	party	value
	chair	hriac	chail	table	total
15	green	ergne	breen	grass	chain
	metal	lmeta	metan	steel	blind
16	water	rtaew	nater	ocean	clerk
	apple	lpepa	plape	fruit	grade
17	groom	omgro	groom	bride	owner
	sweet	wtese	sween	sugar	crime
18	steal	atlse	stean	thief	scout
	wheat	htaew	whean	flour	niece
19	stamp	matps	stame	letter	effort
	store	rteas	storp	market	police
20	nurse	sreun	gurse	doctor	status
	glass	slsag	flass	window	europe