Individual and Developmental Differences in Semantic Priming: Empirical and Computational Support for a Single-Mechanism Account of Lexical Processing

Article *in* Psychological Review · October 2000

DOI: 10.1037/0033-295X.107.4.786 · Source: CiteSeer

2 authors:

David Plaut
Carnegie Mellon University

James Booth
Vanderbilt University

Individual and Developmental Differences in Semantic Priming: Empirical and Computational Support for a Single-Mechanism Account of Lexical Processing

James R. Booth
Department of Psychology
Carnegie Mellon University

David C. Plaut
Departments of Psychology and Computer Science
Center for the Neural Basis of Cognition
Carnegie Mellon University

November 1997
Manuscript submitted for publication. Please do not quote without permission.

Abstract

Existing accounts of single-word semantic priming phenomena incorporate multiple mechanisms, such as spreading activation, expectancy-based processes, and postlexical semantic matching. We provide both empirical and computational support for a single-mechanism, distributed network account. Previous studies have found greater semantic priming (i.e., faster RTs following related vs. unrelated primes) for low- compared with high-frequency target words, and inhibition (i.e., slower RTs following unrelated vs. neutral primes) only at long stimulus-onset asynchronies (SOAs). A series of experiments examined the extent to which these effects depended on individual differences among subjects in age or perceptual ability. Third-grade, sixth-grade, and college students performed a lexical decision task on high- and low-frequency target words preceded by related, unrelated, and nonword primes. We found that greater priming for low-frequency targets was exhibited only by subjects with high perceptual ability, and that this restriction held across differences in age and SOA. We also replicated the finding of inhibition at a long but not short SOA for the college students, but found no inhibition for the children even at the long SOA. We provide an account of these results in terms of the properties of distributed network models, and support this account by demonstrating that an implemented simulation reproduces the empirical findings despite lacking expectancy-based processes and postlexical semantic matching. The results suggest that distributed network models can provide a viable single-mechanism account of lexical processing.

Introduction

One of the most fundamental and important findings in cognitive psychology is that people are faster and more accurate when they read a word such as BUTTER when it is preceded by a related word such as BREAD compared with when it is preceded by an unrelated word such as DOCTOR (Meyer & Schvaneveldt, 1971). This priming effect occurs for word pairs that are either categorically or associatively related (Becker, 1980) and in a variety of tasks, including both word naming and lexical decision (Meyer, Schvaneveldt, & Ruddy, 1975).

The robust nature and generality of priming effects has led theorists to suggest that they reflect fundamental properties of the organization of lexical items within the human cognitive system. Spreading-activation theories (e.g., Anderson, 1983; Collins & Loftus, 1975; McNamara, 1992, 1994; Quillian, 1967) assume that semantic memory consists of a network of interconnected nodes and that activation spreads along the connections in this network. This spread of activation is assumed to be fast and automatic, and therefore causes a related prime to facilitate the processing of a target word (Balota & Chumbley, 1984; Chumbley & Balota, 1984). Compound-cue theories (e.g., Dosher & Rosedale, 1989; Ratcliff & McKoon, 1988) propose that, in processing a word, semantic memory is accessed using a cue consisting of the word conjoined with the context in which it occurs (i.e., the preceding word). Because related words co-occur more frequently than do unrelated words, their compound cues tend to have greater familiarity, resulting in faster retrieval (according to most general memory models; e.g., Gillund & Shiffrin, 1984; Hintzman, 1986; Murdock, 1982).

Distributed network models have also been proposed to account for semantic priming effects (Becker, Moscovitch, Behrmann, & Joordens, 1997; Borowsky & Masson, 1996; Joordens & Becker, 1997; Kawamoto, 1988; Masson, 1991, 1995; McRae, de Sa, & Seidenberg, 1997; Moss, Hare, Day, & Tyler, 1994; Plaut, 1995; Sharkey & Sharkey, 1992). The fundamental assumption in such models is that concepts are represented by distributed patterns of activity over a large number of interconnected processing units, such that related concepts are represented, not by interconnected "localist" concept nodes as in spreading-activation models, but by similar (overlapping) patterns. Semantic priming arises because, in processing a target, the network starts from the pattern produced by the prime, which is more similar to the representation of the target for a related prime compared to an unrelated prime.[1] In some formulations (e.g., Moss et al., 1994; Plaut, 1995), associations among words are reflected independently by an increased probability of the transition from one concept to another during training. Associative priming thus arises because the network has learned to derive the representation of a target word more frequently, and hence more effectively, when starting from the representation of an associated prime word compared with a nonassociated prime.

All of these theoretical frameworks—spreading-activation theories, compound-cue theories, and distributed network theories—are challenged to varying degrees by the considerable body of research showing that priming effects are influenced by a variety of experimental factors, including target frequency, category dominance, relatedness proportion, stimulus quality, stimulus-onset asynchrony (SOA), and the task performed by subjects (see Neely, 1991, for a review). The almost universal response to these challenges is to complicate theories of lexical processing by postulating additional mechanisms that collectively account for the range of findings, albeit in a post hoc manner.

A good example of a multi-mechanism account is Neely and Keefe's (1989) Hybrid Three-Process Theory, in which spreading activation is augmented with expectancy-based processes and with retrospective semantic matching in attempting to explain all of the priming effects in lexical decision and naming. According to this theory, presentation of a prime first engages automatic spreading activation processes. In addition, subjects are assumed to use the prime to generate two expectancy sets of possible targets: a set of visually similar items and a set of semantically similar items. When the target is presented, subjects first search for the target among items in the semantic set in random order; if it is not found, subjects then search through the items in the visual set in order of their frequency. Performance is assumed to be facilitated if the target is found and inhibited if it is not. In addition, following lexical access of the target but before executing a response, subjects retrospectively compare the target with the prime to determine whether they are related; performance is further facilitated if they are and inhibited if they are not.

The central goal of the current work is to provide both empirical and computational support for a more parsimonious account of semantic priming phenomena, and lexical processing more generally, in terms of the properties of distributed network models. The work focuses on two particular sets of findings, concerning 1) the effects of target frequency on the magnitude of semantic priming, and 2) the degree to which priming effects result from facilitation and/or inhibition (relative to a neutral prime baseline) as a function of SOA. In brief, previous studies have found greater priming for low- compared with high-frequency targets (Becker, 1979; Borowsky & Besner, 1993) and inhibition at long but not short SOAs (Becker, 1980; den Heyer, Briand, & Smith, 1985b; Smith, Briand, Klein, & den Heyer, 1987). Accounting for these findings has required the addition of considerable complexity to spreading-activation theories, including the separation of frequency and semantic context effects (Borowsky & Besner, 1993) and the introduction of additional strategic, expectancy-based processes (Becker, 1980; Neely, 1977; Paap, Newsome, McDonald, & Schvaneveldt, 1982).[2] We present the results of empirical studies which examined the extent to which these empirical findings depend on individual differences among subjects in age and in perceptual ability. In brief, we found that greater priming for low-frequency targets was exhibited only by subjects with high perceptual ability, and that this finding held across differences in age and SOA. We also replicated the finding of inhibition at a long but not short SOA for adults, but found no inhibition for children even at the long SOA. We provide an account of these results in terms of the properties of distributed network models, and support this account by demonstrating that an implemented simulation that does not separate frequency and context effects, and which lacks expectancy-based processes, nonetheless reproduces the relevant empirical findings. We consider the strengths and limitations of the approach, and how it might be extended to account for additional semantic priming phenomena, in the General Discussion.

Effects of Target Frequency on Semantic Priming

The finding of an interaction of target frequency and priming context in lexical decision tasks, such that priming effects are larger for low- compared with high-frequency targets, has important implications for theories of lexical processing. This frequency-by-context interaction has been demonstrated using both sentence contexts (e.g.,

[1] Although, in some sense, activation "spreads" within a distributed network, this spread is not between concepts but between features of concepts, sometimes termed *microfeatures* (Hinton, McClelland, & Rumelhart, 1986). When a given concept is "active"—that is to say, its pattern is present over the units—all other concepts are simultaneously active to the degree that their patterns overlap with that of the "active" concept.

[2] We will contrast our distributed network account primarily with spreading-activation accounts because such accounts have been elaborated in the greatest detail to address the issues under consideration. We believe, however, that similar issues arise with respect to compound-cue theories (e.g., Ratcliff & McKoon, 1988).

Stanovich & West, 1981; Stanovich, West, & Feeman, 1981) and single-word prime (Becker, 1979; Borowsky & Besner, 1993). The traditional account of this interaction within a spreading-activation framework is an extension of the logogen model of word recognition (Morton, 1969). This model assumes that the resting level of activation for a word detector is further from threshold for low-frequency words than for high-frequency words, resulting in a larger effect of priming context on the former than on the latter. A potential problem with this explanation, however, is that priming context also interacts with stimulus quality, but target frequency and stimulus quality do not interact with each other (Borowsky & Besner, 1993). This pattern of results makes it difficult—at least within an additive factors framework (Sternberg, 1969)—to locate context and frequency effects at the same stage of processing. This led Borowsky and Besner (1993) to postulate that, whereas priming effects arise within semantics, frequency effects are due to the mapping between orthography and semantics. On their account, this mapping is stronger for high-frequency words because they have been encountered more often during reading than low-frequency words. Consequently, high-frequency target words generate more rapid activation of semantics from orthography and, therefore, are less affected by priming context, than are low-frequency targets.

Thus, spreading activation theories can account for the interaction of priming context and target frequency, if it is assumed that these factors influence different stages of processing. However, certain other findings in the priming literature—the relative effects of facilitation and inhibition as a function of SOA in skilled readers, and individual differences in priming effects as a function of age and/or reading ability—have prompted the introduction of additional complexities into spreading-activation theories.

Effects of Stimulus-Onset Asynchrony on Facilitation and Inhibition in Priming

Much of the research on semantic priming in adults has focused on the time course of facilitation and inhibition in naming and lexical decision tasks. In these studies, *facilitation* is defined as a decrease in reaction time (RT) to a target word following a related priming context compared to a neutral context (e.g., a nonword or a string of Xs), whereas *inhibition* is defined as an increase in RT following an unrelated context versus a neutral context. In general, priming effects are smaller at short SOAs (i.e., less than 250 ms) compared with long SOAs ($>$ 800 ms). Furthermore, the effects at short SOAs are due only to facilitation, for both categorical and associative priming. At long SOAs, associative priming effects still result primarily from facilitation, whereas categorical priming effects result from both facilitation and inhibition (Becker, 1980; den Heyer et al., 1985b; Smith et al., 1987).

These findings have generally been interpreted in terms of a distinction between automatic and strategic processes. Posner and Snyder (1975) argued that all cognitive processes can be characterized by the two different mechanisms of spreading activation and strategic attention. Spreading activation is automatic and occurs without intention, whereas strategic processes require conscious attention and are of limited capacity. In an extension of this dichotomy, Neely (1977) suggested that facilitation in word recognition results from fast, automatic spreading activation, whereas inhibition results from conscious, expectancy-based processes. Expectancy-based processes are slow and strategic because they involve the explicit generation of a set of potential targets from the prime. Processing is assumed to be facilitated if the set contains the actual target and inhibited if it does not (Becker, 1980; also recall Neely & Keefe's, 1989, hybrid three-process model).

We will refer to models that postulate separate spreading-activation and expectancy-based processes as *dual-mechanism* models. Such models account for the aforementioned findings concerning the relative time course of facilitation and inhibition in the following way. At a short SOA, the recognition of a target can be facilitated by a related prime as a result of fast, automatic spreading activation, but it cannot be inhibited by an unrelated prime because there is insufficient time for the slow, expectancy-based processes to operate. At long SOAs, expectancy-based processes (along with spreading activation, in some formulations) have time to influence word recognition, so the priming effects result from both facilitation and inhibition (see Neely, 1991).

Developmental and Individual Differences in Priming

Additional constraints on theories of semantic priming come from studies of developmental differences in the influence of priming context on word recognition. A number of studies have shown larger priming effects for younger and poor readers than for older and good readers when reading target words presented after a single-word or sentential priming context (e.g., Schwantes, 1985, 1991; Simpson & Lorsbach, 1983, 1987; West & Stanovich, 1978). Other studies have investigated contextual processes in children by examining their oral reading errors (Biemiller, 1970; Goldsmith-Phillips, 1989; Jackson & Biemiller, 1985; Wijnen, 1992). These studies show that the oral reading errors of older and good readers tend to be phonemically related to the text being read, whereas the errors of younger and poor readers tend to be semantically or syntactically related. Taken together, studies measuring reaction times and oral reading errors establish that younger and poor readers show larger contextual effects than older and good readers.

The most commonly cited account of the developmental differences in word priming effects is Stanovich's (1980) *interactive compensatory* model. This dual-mechanism model was influenced by Perfetti and Lesgold's (1977) verbal encoding model which argued that "contextual processes are limited by word coding processes" (p. 273). According to these models, older and good readers have fast and automatic word decoding skills and, thus, rely less on expectancy-based processes to facilitate or inhibit word recognition. By contrast, because their word decoding is slower and less automatic, younger and poor readers rely more heavily on expectancy-based processes (Raduege & Schwantes, 1987) and, thus, are expected to exhibit a greater degree of inhibition.

With regard to effects of target frequency, a further relevant property of the interactive compensatory model is that higher-level operations (e.g., expectancy-based processes) affect lower-level processes (e.g., word recognition) only when the latter are slow or strategic. It is well known that, for adults, word recognition is faster for high- compared with low-frequency words (see Monsell, 1991, for review). For children, however, the decoding of most high-frequency words is not automatized to adult levels until the middle elementary school years (see, e.g., Golinkoff & Rosinki, 1976; Guttentag & Haith, 1979; Perfetti, Finger, & Hogaboam, 1978; Perfetti & Hogaboam, 1975). Children should, therefore, use priming context to an equal degree for recognizing high- and low-frequency words because they decode all words slowly. In this way, the interactive compensatory model predicts a three-way interaction of age or ability, priming context, and target frequency. Specifically, older and high-ability readers should show greater priming effects for low- compared with high-frequency target words, whereas younger and low-ability readers should show equivalent priming for these items.[3] Moreover, the model predicts that older and high-ability readers should show both facilitation and inhibition for low-frequency targets but only facilitation for high-frequency targets, whereas the younger and low-ability readers should show both facilitation and inhibition for both low- and high-frequency targets.

The existing literature on the relative contribution of facilitation and inhibition to priming context effects for younger and low-ability readers is inconclusive. Most studies on developmental differences in priming effects have employed entire sentences as context manipulations, and have found that older children and adults exhibit both less facilitation and less inhibition compared with younger children (Schwantes, Boesl, & Ritz, 1980; Stanovich, Nathan, West, & Vala-Rossi, 1985; West & Stanovich, 1978). However, these sentence priming paradigms differ in important ways from single-word priming paradigms. Sentence priming is influenced by syntactic and discourse-level factors, such that the developmental decrease in inhibition from sentence contexts may be due to greater efficiency of sentence integration processes with increased reading experience (Simpson & Lorsbach, 1987). By contrast, single-word priming is assumed to reflect lexical processing more purely, and may give rise to a different developmental pattern of facilitation and inhibition.

To our knowledge, only two studies have examined developmental differences in facilitation and inhibition using the single-word priming context paradigm with a neutral prime baseline (i.e., a string of Xs). In a naming study involving second-grade, fourth-grade, sixth-grade, and college students, Simpson and Lorsbach (1983) found that younger students exhibited more facilitation than older students at low and high relatedness proportions (25% and 75%), but less inhibition at high relatedness proportions. In a second naming study with fourth-grade and sixth-grade students, Simpson and Lorsbach (1987) found that good readers showed more inhibition but less facilitation than poor readers at high relatedness proportions (75%). These results suggest that, contrary to the predictions of the interactive compensatory model, facilitation at the lexical level decreases with development, whereas inhibition increases with development. This conclusion is also supported by research indicating that the ability to use inhibitory processes in picture naming and Stroop tasks is still developing during the early elementary school years (Schadler & Thissen, 1981; Guttentag & Haith, 1979).

As far as we know, no empirical studies of developmental or individual differences have investigated whether the magnitude of the interaction between a single-word priming context and target frequency is greater for older and high-ability readers than for younger and low-ability readers, and no developmental studies have investigated the time course of facilitation and inhibition using nonword primes as a neutral baseline condition. Our empirical and computational modeling work did exactly this.

Most studies of individual differences in semantic priming effects have focused on the impact of overall reading skill, as indexed by standardized tests of naming accuracy and/or reading comprehension. Few studies have examined the relative contribution of specific aspects of reading skill. Thus, we do not know whether the differences in semantic priming result from variations in higher-level reading skills, such as vocabulary knowledge and inferential processes, or in lower-level reading

[3] There is evidence for a two-way interaction between reading skill and target frequency, such that the difference in RTs for low- and high-frequency words is greater for low- compared with high-ability readers. For example, Perfetti and Hogaboam (1975) found that third- and fifth-grade low-ability readers named high-frequency words about 200 ms slower but named low-frequency words about 1000 ms slower than did their high-ability counterparts (also see Schwantes, 1991).

skills, such as perceptual encoding ability. While vocabulary knowledge is a strong determinant of reading ability (Stahl, Hare, Sinatra, & Gregory, 1991), perceptual efficiency, as measured by match-to-sample tasks, is also related to reading proficiency (Vernon, 1987). Moreover, deficits in rapid perceptual processing are strongly associated with abnormal reading acquisition (Booth, Hunt, Perfetti, & MacWhinney, 1997; Eden, Stein, Wood, & Wood, 1995; Eden, VanMeter, Rumsey, Maisog, Woods, & Zeffiro, 1996; Lovegrove, Martin, & Slaghuis, 1986) and abnormal language development more generally (Tallal, Miller, & Fitch, 1993; Tallal & Piercy, 1973, 1974, 1975, see Farmer & Klein, 1995, for review). In fact, perceptual efficiency may be particularly relevant in the early stages of reading acquisition. Detterman and Daniel (1991) found that the correlation of perceptual efficiency measures with the Wechsler IQ score was $r = .26$ for high-IQ subjects but $r = .60$ for low-IQ subjects. Given the typically strong relationship found between IQ and reading skill, the latter high correlation suggests that lower-level perceptual abilities play an important role in the development of reading skill. Our empirical studies considered directly whether individual differences in perceptual ability, as measured by a match-to-sample task, has an important impact on the use of priming context for recognizing high- versus low-frequency target words.

Distributed Network Models of Semantic Priming

A spreading-activation framework can account for the interaction of priming context with target frequency, as well as the relative time course of facilitation and inhibition, but not without being elaborated to include discrete processing stages and separate, expectancy-based processes. Our central theoretical claim is that distributed network models can also account for these empirical findings, as well as the novel ones reported below, without these added complexities, and hence provide a more parsimonious explanation than do spreading-activation theories. The approach has the added benefit of being supported by computational simulations that make fully explicit the underlying mechanism that actually gives rise to the appropriate effects.

Our account takes as its starting point a preliminary distributed network simulation developed by Plaut (1995). Although the simulation was not applied to modeling specific empirical data, it exhibited a number of effects that are relevant in the current context. The network was trained on an abstract version of the task of mapping from written words to their meanings. The written form of each word was represented by a particular pattern of activity over a set of orthographic units, while its meaning was represented by another pattern over a set of semantic units. Semantic relatedness among words was encoded by the degree of feature overlap in their semantic representations, whereas associative relatedness was encoded by the frequency with which one word followed another during training (also see Moss et al., 1994). Words also differed in their frequency of presentation during training and in their degree of category dominance (i.e., how similar they were to a prototype pattern for their category), and target words were presented both at normal and reduced contrast (i.e., less binary orthographic input). When tested for priming effects following training, the network exhibited greater associative priming for low-frequency targets than for high-frequency targets, as has been found with human subjects (Becker, 1979; Borowsky & Besner, 1993).[4]

The basis of the frequency-by-context interaction in the network's performance can be understood in terms of the nonlinear effects of the S-shaped sigmoid activation function that relates the input of each unit to its activation (see Figure 1). The use of a sigmoid or logistic function for units is standard in connectionist modeling and can be understood as optimal for tasks involving binary output patterns (see Rumelhart, Durbin, Golden, & Chauvin, 1995).

In processing a target word, the RT of the network is taken to be the point at which the activations of all semantic units approach asymptote (either 0 or 1) and stop changing. This, in turn, depends on the magnitude of the input to the unit—stronger input (positive or negative) drives a unit to asymptote more quickly than weaker input.[5] One factor that influences the strength of the input to the semantic units for a given target word is the frequency with which the word was presented during training. In general, high-frequency words generate stronger input to the semantic units than do low-frequency words because, by being trained more often, they have a greater impact on the weights learned by the network. All else being equal, this stronger input causes the network to settle faster, and thus respond more quickly, to high- compared with low-frequency targets. Another relevant factor is associative relatedness. The network was trained to process the target when starting from the representation of an associated prime more often than when starting from the represen-

[4] In addition to target frequency, stimulus quality also interacted with priming context (i.e., greater priming for degraded compared with intact stimuli) but stimulus quality did not interact with target frequency. Thus, the Plaut (1995) network exhibited the pattern of results found empirically by Borowsky and Besner (1993) and taken by them to imply that frequency and context effects must be located at distinct processing stages. Note, however, that frequency and context effects are not restricted to specific stages or levels of representation within the network; rather, these factors influence weight changes throughout the network over the course of learning.

[5] Strictly speaking, the magnitude of the input to a semantic unit controls not only the time to reach asymptote but also the level of activation reached. We assume that the time required for interactions with other units to push the unit to an extreme activation value (1 or 0) depends primarily on the initial activation that would be produced by the strength of "bottom-up" input, as depicted in Figure 1.

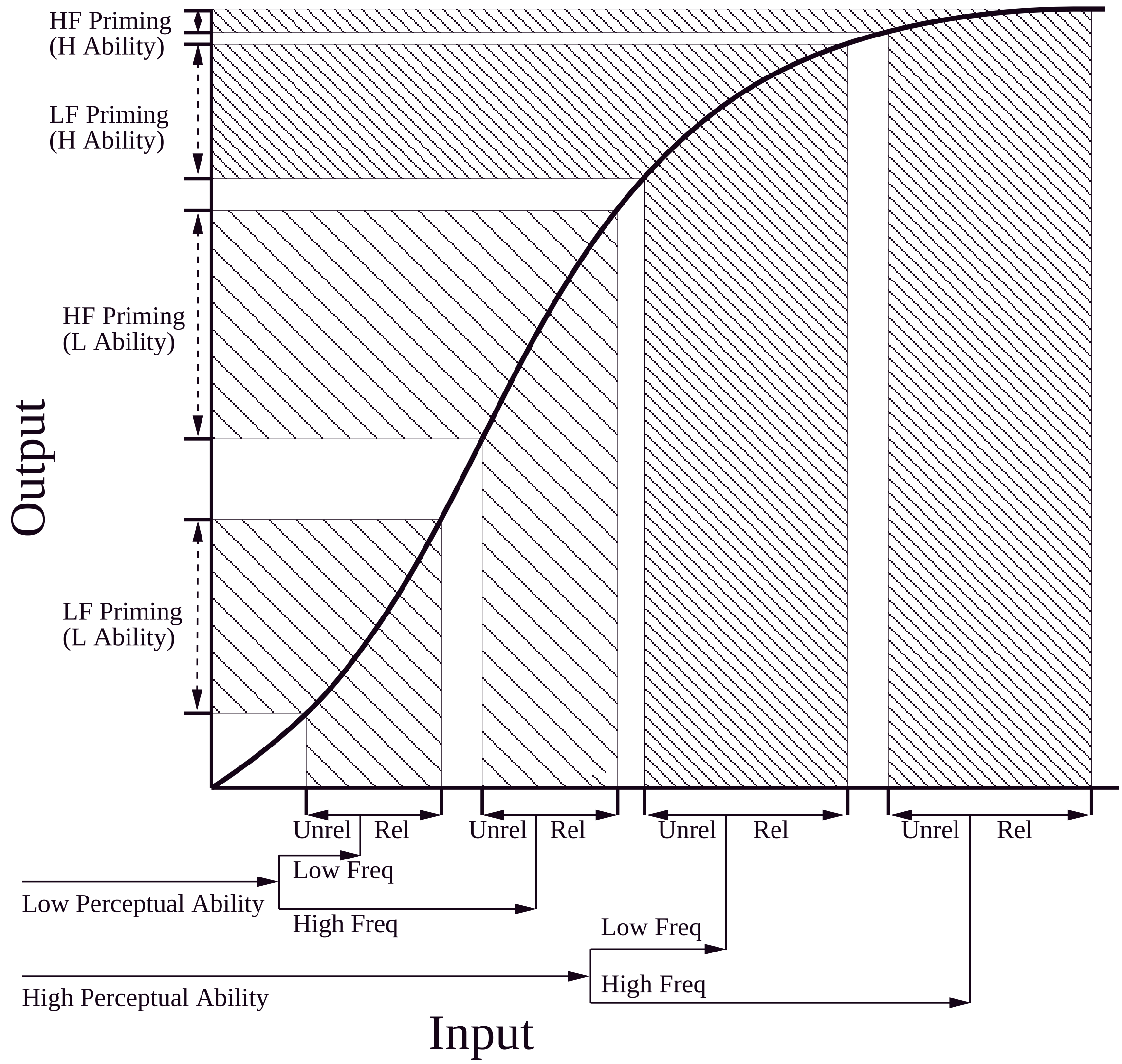

Figure 1. A depiction of how nonlinearities in the sigmoid activation function for semantic units in a distributed attractor network can give rise to greater priming [i.e., the difference in performance following related ("Rel") versus unrelated ("Unrel") primes] of low- versus high-frequency target words ("LF" and "HF", respectively) for subjects with high perceptual ability ("H Ability", narrow-lined regions), but approximately equal priming for low- and high-frequency words for subjects with low perceptual ability ("L Ability", wide-lined regions). The combination of arrows at the bottom depicts the separate contributions of perceptual ability, target frequency, and priming context which are summed together to form the input to a given semantic unit (indicated by the small vertical lines on the x-axis), to which the sigmoid function is applied to determine the activation of the unit. Note that relative magnitudes of these contributions are assumed to be greater for high- compared with low-ability subjects, for high- compared with low-frequency targets, and positive for related primes but negative for unrelated primes (reflecting both facilitation and inhibition, respectively). Also, the magnitudes of the contributions due to target frequency and priming context are assumed to be greater for high-ability subjects because they can process both primes and targets more effectively than low-ability subjects. The bottom portion of the sigmoid function is omitted for clarity.

tation of any particular nonassociated prime, resulting in stronger input to semantic units, and thus faster responses, in a related versus unrelated priming context.

The magnitude of the effect of priming context is, however, modulated by target frequency (see the narrow-lined regions in Figure 1). Specifically, high-frequency words provide sufficient input to the semantic units to boost their activation near asymptote, leaving little room for priming context to have an additional effect on the units' output. By contrast, the input for low-frequency words remains closer to the linear region of the activation function, where further differences due to priming context are reflected more directly in the output of units. Thus, frequency and context interact in the activation of semantic units and, hence, in the settling time of the network, due to the "diminishing returns" of one factor when another factor is sufficiently strong on its own.[6]

With regard to developmental and individual differences, the current formulation of a distributed network account makes a similar prediction as the interactive compensatory model (Stanovich, 1980)—namely, that the frequency-by-context interaction should hold only for older and high-ability readers; younger and low-ability readers should show equal priming effects for high- and low-frequency words. This follows under the assumption that the overall strength of input is weaker for younger and low-ability readers compared with older and high-ability readers, such that both low- and high-frequency words fall within the linear range of the activation function (see the wide-lined regions in Figure 1).[7] In this case, the relative effects of related versus unrelated primes will be roughly equivalent for both low- and high-frequency words. Thus, both the distributed network model and the interactive compensatory model predict that there should be no interaction between priming context and target frequency for low-ability readers, but that there should be an interaction between priming context and target frequency for high-ability readers.

There are fundamental differences, however, in how these two models view the changes in performance between children and adults. An important implication of our single-mechanism account of semantic priming is that children and adults should differ only quantitatively. For example, because children have less reading experience than adults, they benefit from fewer learning episodes with any given word (i.e., their effective word frequencies are lower) and, hence, they are slower at processing a prime. Consequently, a longer SOA in children may result in the same degree of activation in the semantic system as a shorter SOA in adults. Thus, on a distributed network account, priming effects in children at a long SOA would be expected to be similar to priming effects in adults at a short SOA. By contrast, a dual-mechanism model like the interactive compensatory model (Stanovich, 1980) holds that children are qualitatively different from adults because priming effects in children result more from inhibitory expectancy-based processes, whereas priming effects in adults result more from spreading activation processes.

Finally, a distributed network approach may also be able to account for the occurrence and time course of facilitation and inhibition in priming, without invoking separate expectancy-based processes. As mentioned earlier, associative priming seems to result from facilitation at both short and long SOAs, whereas categorical priming seems to result from facilitation at short SOAs but from both facilitation and inhibition at long SOAs. In the Plaut (1995) model, associative priming is due to the increased frequency with which targets are preceded by associated versus nonassociated primes during training. Plaut showed that associative priming effects in the increase as the duration of the prime is lengthened, because the resulting pattern more closely approximates the representation of the prime that is associated with that of the target. By contrast, categorical priming effects, which are due to semantic feature overlap among category members, peak at a relatively short prime duration and then decrease with additional processing of the prime. This decrease is due to the fact that semantic units, including those that *differ* between the prime and target, are being driven towards their asymptotic values. These differences take time to be reversed when the target is presented, thereby diminishing the advantage of starting with some overlapping features due to a categorically related prime.

In other words, at longer SOAs, the network exhibits a greater degree of *hysteresis* in moving from one stable state to another, even when those states share many active units. It is important to understand the basis for this effect because, as we will argue later, it may explain the shift from facilitation dominance at short SOAs to inhibition dominance at long SOAs (Becker, 1980). Hysteresis in the network arises from the operation of *attractors* over semantic representations. During the course of training, the network learned to make the semantic representation of each word into an attractor, such that similar but unfamiliar semantic patterns gradually settle to the near-

[6] Directly analogous explanations based on the nonlinearity of the sigmoid function have been proposed by Cohen, Dunbar, and McClelland (1990) to account for the balance of facilitation and inhibition in Stroop tasks, and by Plaut, McClelland, Seidenberg, and Patterson (1996) to account for the interaction in naming latencies of frequency and spelling-sound consistency (e.g., Seidenberg, Waters, Barnes, & Tanenhaus, 1984a; Taraban & McClelland, 1987; Waters & Seidenberg, 1985), and for the three-way interaction of frequency, consistency, and imageability (Strain, Patterson, & Seidenberg, 1995).

[7] The assumption that good and poor readers differ only in how strongly semantics is activated by orthography is clearly a simplification. It is likely that good and poor readers differ in many ways, including the quality of their semantic representations, but this is not the focus of our empirical or modeling work.

est familiar (word) representation. Attractor patterns can be thought of as particular points in a high-dimensional *state space* containing a dimension for each semantic unit. Within this space, the settling process consists of movement of the point corresponding to the initial pattern in the network downhill into a bowl-like *basin of attraction* to the attractor point at the bottom of the basin. When a target word follows a prime, the network must alter its activity pattern to move up and out of the attractor basin for the prime in order to settle to the bottom of the basin for the target. To the extent that the processing of the prime is particularly strong or prolonged (e.g., at a long SOA), the network will require more time to move out of its basin of attraction, resulting in a slower RT to the target compared with when a prime (e.g., a nonword) produces only weak activation (e.g., at a short SOA). Note that associative priming, by contrast, is not subject to this hysteresis because the network learned to make associated prime-target transitions effectively due to their elevated frequency during training.

Although the pattern of results exhibited by the Plaut (1995) network is broadly consistent with the empirical findings, the relative magnitudes of facilitation and inhibition were not established explicitly by comparing RTs to target words following neutral primes (e.g., nonwords or Xs) to those following related and unrelated primes. In fact, as far as we know, the computational model of semantic priming presented below is the first to be tested directly in this manner.

In summary, distributed network models suggest an account for individual and developmental differences in the interaction between priming context and target frequency. These models are in stark contrast to dual-mechanism models which postulate a strategic, expectancy-based mechanism separate from an automatic, spreading-activation mechanism. In the current work, we examine the degree to which the interaction between single-word priming context and target frequency depends on individual and developmental differences at long and short SOAs, and we present a distributed network model that accounts for the findings and makes additional empirical predictions. Specifically, three empirical studies were conducted to determine whether age and perceptual ability modulated the typical interaction found in adults between priming context and target frequency (Becker, 1979; Borowsky & Besner, 1993). The first experiment examined differences in perceptual ability among college students at a long SOA (800 ms). The second experiment investigated ability and age differences among third- and sixth-grade children at the same long SOA. The third experiment examined ability differences among college students at a short SOA (200 ms). We predicted that older and high-ability readers would exhibit an interaction between priming context and target frequency, whereas younger and low-ability readers would not. However, we expected these priming effects to be composed of different patterns of facilitation and inhibition for the adults versus the children. For the adults, based on previous results (Becker, 1980; den Heyer et al., 1985b; Smith et al., 1987), the context-by-frequency interaction should result only from facilitation at the short SOA but from both facilitation and inhibition at the long SOA. For the children at the long SOA, the interaction between context and frequency should result primarily from facilitation according to the distributed network account, but from inhibition according to the interactive compensatory account.

Following the empirical experiments, a computational simulation of a distributed network model is presented. Following Plaut (1995), a network was trained to map orthographic representations of words onto semantic representations, including both associative relatedness (increased transition probabilities) and semantic relatedness (increased semantic feature overlap). Individual differences in perceptual ability were implemented by manipulating the strength of the orthographic input, and developmental differences were modeled by examining the performance of the network at different points in training. Word frequency was reflected in the frequency with which words were presented during training. Finally, SOA corresponded directly to the timing of prime and target presentation during testing. Simulation results support our claim that distributed network models offer a viable alternative to dual-mechanism models of semantic priming.

Experiment 1

The primary purpose of Experiment 1, with a long SOA (800 ms), was to test our prediction that college students with high perceptual ability should use priming context more for the recognition of low-frequency words than for the recognition of high-frequency words, whereas college students with low perceptual ability should use priming context equally for facilitating the recognition of high- and low-frequency words. Given the long SOA, this interaction was also expected to result from both facilitation and inhibition.

Method

Subjects. Subjects were 94 college students (*M* age = 20.9; *SD* = 5.7) at the University of Maryland who participated to fulfill a psychology course requirement. All subjects had English as a first language and reported that their vision was corrected to normal.

Apparatus. Subjects viewed all stimuli for the priming task on a VGA monitor controlled by an IBM 286 computer with Micro Experimental Laboratories (MEL) software (Schneider, 1990). The subjects controlled stimulus presentation and recorded their responses with a computer keyboard. MEL computed RTs by measuring the time lapse between the target word and the subject's response. Error rates were also recorded by MEL.

Materials and Design. The critical stimuli for the priming task, listed in Appendix 1, were 120 prime-target pairs in each of three conditions: unrelated word prime and word target (e.g., EIGHT-BELOW), related word prime and word target (e.g., ABOVE-BELOW), and nonword prime and word target (e.g., KARBS-BELOW). Each type of prime-target condition had an equal probability of being presented to each subject (i.e., 40 trials). The nonword target pairs were 40 different word primes (e.g., HAPPY-GORPH) and 40 different nonword primes (e.g., ZENOX-AJUPE) paired with 80 different nonword targets.[8] These 5 conditions totaled 200 test pairs. Note that the word and nonword stimuli were not matched orthographically. For example, the nonwords have lower summed positional bigram and trigram frequencies (based on the Kučera & Francis, 1967, corpus) than the words [bigrams: *M*s = 62.6 vs. 82.0, respectively, $F(1,518) = 37.86$, $MSE = 1106$, $p < .001$; trigrams: *M*s = 6.3 vs. 11.9, respectively, $F(1,518) = 74.50$, $MSE = 46.4$, $p < .001$]. This difference will be relevant to the design of the stimuli used in the computational simulation.

Each prime was presented in white lowercase letters on a black background, and was followed by the target words presented in lowercase letters.[9] Targets were presented at an 800 ms SOA with a 200 ms inter-stimulus interval (ISI). The three conditions for the critical prime-target pairs were counterbalanced between-subjects. Specifically, a related prime, a nonword prime, and an unrelated prime preceded the same target word equally often across three different experimental lists. Because three counterbalancing lists were used, the same stimulus item was never seen by a single subject on more than one occasion. Within each list, the order of item presentation was randomized within-subjects. There were also 30 practice trials that consisted of 6 pairs of the 5 prime-target conditions. The practice trials were excluded from all statistical analyses.

The strength of association between the related prime and target word ($M = .47$, $SD = .17$) was controlled for by using established association norms (Nelson, McEvoy, & Schreiber, 1994), because controlling for this association enables a better understanding of priming effects on word recognition (Becker, 1980). All stimuli were also restricted to be five letters in length because demonstrating interactions of target frequency with priming context is more convincing if confounding factors such as word length are controlled for. In addition, target frequencies were chosen such that they were normally distributed after transforming frequency using the formula $40 + \log_{10}(f + 1)$, where f is the frequency of the target in Kučera and Francis (1967). This logarithmic transformation reduced the very large variability typical of target frequency counts (see Borowsky & Besner, 1993). Frequency was also dichotomized into high frequency ($M = 232.6$; $SD = 167.7$) and low frequency ($M = 30.7$; $SD = 20.5$) for ease of interpretation in the figures; however, all ANCOVAs treated frequency as a continuous variable.

Nonword primes were used as neutral primes instead of repetitive stimuli like the word READY or a string of Xs because these latter stimuli may not engage the linguistic substrates involved in word recognition and they may also lose their alerting qualities over repeated presentations (Antos, 1979; Jonides & Mack, 1984). RTs to target words following these repetitive primes may therefore be inflated resulting in an underestimation of inhibition by unrelated primes and an overestimation of facilitation by related primes (see Neely, 1991, pp. 278–281, for a discussion of using nonwords as neutral primes). Note, however, that recently there has been controversy over whether nonword primes are, in fact, good neutral baselines (see, e.g., McNamara, 1994; McKoon & Ratcliff, 1992).

The Symbol Search Test of the Wechsler Intelligence Scale for Children—Third Edition (WISC-III; Wechsler, 1991) was used as a measure of perceptual processing ability. The Symbol Search Test loads on the Processing Speed factor of the WISC-III. This paper and pencil task required each subject to correctly indicate as fast as and with as few errors as possible whether either of 2 meaningless symbols to the left appeared in a row of 5 meaningless symbols to the right. This test consisted of 3 pages of 15 items each, and each page was timed from the moment in which the subjects checked the first yes or no box to the moment in which they made their last response on that page. A speed/accuracy score was calculated for each subject by dividing their time needed to complete the test by their accuracy. All mean accuracy scores in Experiments 1–3 were more than 43 out of 45.

[8] It should be pointed out that the likelihood of a word target is 0.67 following word primes but only 0.5 following nonword primes. Consequently, subjects could potentially derive and use information about the lexicality of the prime to bias lexical decisions to the target. The effect of this would presumably be to facilitate word responses following word primes (related or unrelated) relative to word responses following nonword primes, thereby increasing the magnitude of facilitation relative to inhibition. This potential bias should be kept in mind when evaluating the patterns of facilitation and inhibition exhibited by subjects in the current and subsequent experiments.

[9] The target words were low-intensity in half of the trials and high-intensity in the other half. The low-intensity targets were dark gray letters presented on a black background, and the high-intensity targets were white letters presented on a black background. At each level of intensity, half of the targets were high-frequency and half were low-frequency. The target words in the related, unrelated, and nonword priming conditions were also counterbalanced between subjects for low- and high-intensity. It turned out that the intensity manipulation produced no reliable effects in any of the current experiments, so all analyses were collapsed across target intensity.

In addition, as a measure of vocabulary knowledge, the subjects were given the Peabody Picture Vocabulary Test (PPVT-R). The PPVT-R was used because of the large vocabulary knowledge range anticipated. The PPVT-R has been normed for 2- to 33-year-olds (Dunn & Dunn, 1981). The PPVT-R is correlated (rs $> .60$) with other standardized vocabulary and verbal measures such as the WISC and the Wechsler Adult Intelligence Scale (WAIS).

The Symbol Search Test and PPVT were chosen specifically to not tap reading processes directly. We assume that performance on the Symbol Search Test reflects the speed with which letters and words can be encoded by the orthographic system, but measured in a way that avoids confounds from other orthographic, lexical, or semantic factors. The PPVT was administered in order to control for higher-level semantic and vocabulary knowledge.

Procedure. All subjects were individually administered the Symbol Search Test, the PPVT-R, and the priming task in a room that was separated by two doors from any sound interference. The Symbol Search Test and the PPVT-R were always administered first. The testing period lasted approximately 60 minutes.

The priming task began with the experimenter reading instructions which were presented on a computer monitor placed about 50 cm in front of the subject. At this distance, the 5-letter target words subtended about 1.5 degrees of visual angle. The experimental session proceeded as follows. The subjects were told that the first stimulus would be either a word or a nonword, and that the second stimulus would also be a word or a nonword. The subjects were then told to decide whether the second stimulus spells a word they know, and to respond as accurately and quickly as possible by pressing the red key (z) on the key board with their left hand if the stimulus was not a word and by pressing the green key (/) with their right hand if the stimulus was a word. The subjects were then told that they could control the rate at which each trial would be presented. Pressing the space bar would make the "Get Ready" indicator disappear and start the trial by causing a fixation cross (+) to be presented on the screen. The subjects were asked to fixate on this cross and were told that after 2 seconds the first stimulus would appear for less than a second, and then the second stimulus would appear shortly thereafter. The subjects were told that the second stimulus would remain on the screen until they responded. There was then a mandatory 2 second inter-trial interval.

Results and Discussion

Subjects were dichotomized into high- or low-perceptual-ability groups based on their Symbol Search Test speed-accuracy score (see Table 1). The low-ability group scored significantly poorer than the high-ability group, $t(92) = 13.34$, $p < .05$. The vocabulary measure (PPVT-R) did not correlate significantly with the perceptual measure suggesting that these instruments were measuring two distinct underlying abilities ($r = .09$). This independence was supported further by the finding of no significant differences in vocabulary knowledge between the high- and low-perceptual-ability groups ($t < 1$). Therefore, any differences in priming between the high- and low-perceptual-ability groups could not be due to vocabulary differences.

In all subsequent ANCOVAs, perceptual ability (high vs. low) was a between-subject factor, priming context (related vs. unrelated) was a within-item factor, and target frequency was a continuous between-item factor. Target frequency was treated as a continuous factor to more accurately reflect the underlying variable of frequency, although it meant that only item analyses could be computed. Note, however, that all figures treat target frequency as a dichotomous variable (high vs. low) for ease of interpretation. If the three-way interaction between perceptual ability, priming context, and target frequency was significant, separate ANCOVAs were computed for the high- and low-ability groups to determine whether priming context interacted with target frequency for that ability level. Planned t-tests were then calculated to compare the related, unrelated, and nonword priming conditions in order to determine whether priming effects resulted from facilitation, inhibition, or both.

For the semantic priming task, an item with a correct mean RT of greater than 2.5 *SD*s from the overall mean in the related, unrelated, or nonword prime condition for either the high- or low-ability subjects was eliminated (3 items or 2.5%). Correct mean RTs or error rates for each word were then entered into a 2 (perceptual ability: high, low) $\times$ 2 (priming context: related, unrelated) ANCOVA with target frequency as a continuous regressor variable. The mean RTs are shown in Figure 2. The RT analysis yielded a main effect of perceptual ability, $F(1,468) = 227.50$, $MSE = 811{,}834$, $p < .001$, priming context, $F(1,468) = 12.02$, $MSE = 42{,}897$, $p < .001$, and target frequency, $F(1,468) = 17.94$, $MSE = 64{,}031$, $p < .001$. However, these main effects were qualified by a significant three-way interaction between perceptual ability, priming context, and target frequency, $F(1,468) = 4.47$, $MSE = 15{,}947$, $p < .05$.

The interaction between priming context and target frequency was then analyzed separately for the high- and low-ability subjects. For the high-ability subjects, there were significant main effects of priming context, $F(1,234) = 8.18$, $MSE = 23{,}160$, $p < .01$, and target frequency, $F(1,234) = 5.47$, $MSE = 15{,}492$, $p < .01$. For the low-ability subjects, there were also significant main effects of priming context, $F(1,234) = 4.60$, $MSE = 19{,}803$, $p < .01$, and target frequency, $F(1,234) = 12.65$, $MSE = 54{,}471$, $p < .001$. As predicted, there was a significant interaction between priming context and target frequency for the

Table 1

Means (and Standard Deviations) for Age and for the Perceptual and Vocabulary Ability Measures for Experiment 1

	Age (in years)	Perceptual Speed-Accuracy	Vocabulary Raw Score
All Subjects (*N* = 94)	20.9 (5.7)	2.5 (0.4)	159.7 (9.3)
High Perceptual Ability (*N* = 47)	20.1 (4.3)	2.2 (0.2)	159.2 (9.7)
Low Perceptual Ability (*N* = 47)	21.9 (6.8)	2.8 (0.3)	159.7 (9.1)

Note. Perceptual Speed-Accuracy is mean time (in seconds) divided by number correct (out of 45) on the Symbol Search Test of the WISC-III (Wechsler, 1991). Vocabulary Raw Score is the mean raw score (out of 171) on the PPVT-R (Dunn & Dunn, 1981).

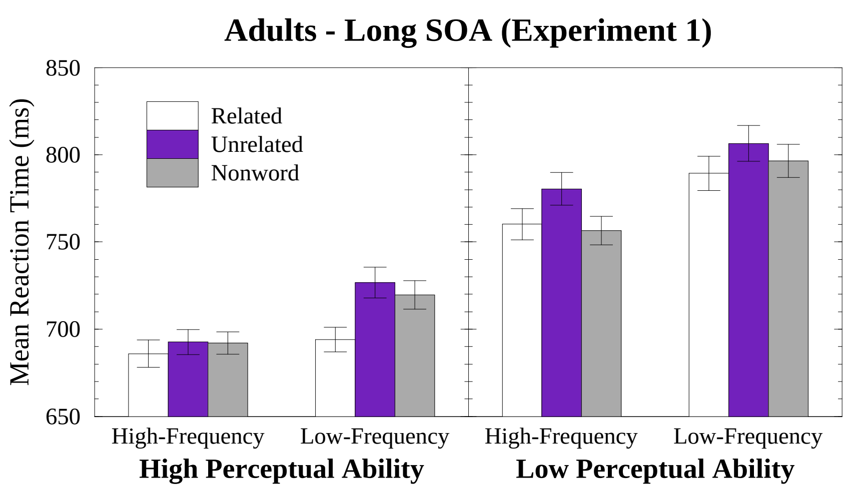

Figure 2. Item means for Experiment 1 of correct mean RTs to high- and low-frequency target words following related, unrelated, and nonword primes (800 ms SOA), for high- and low-perceptual-ability college students. Error bars indicate 1 standard error.

high-ability subjects, $F(1{,}234) = 6.07$, $MSE = 17{,}173$, $p < .05$, but not for the low-ability subjects ($F < 1$). For high-ability subjects, priming context influenced the recognition of low-frequency targets (priming d = 33 ms) more than the recognition of high-frequency targets (d = 7 ms). By contrast, for low-ability subjects, priming context influenced the recognition of high- and low-frequency targets to a similar degree (ds = 20 and 17 ms, respectively).

As suggested in the Introduction, the observed interaction among perceptual ability, priming context, and target frequency can be understood in terms of the asymptotic response of semantic units within a distributed network model (see Figure 1). High-perceptual-ability subjects have efficient perceptual processes which cause the orthographic input for high- but not low-frequency targets to strongly activate their semantic features, thereby leaving little room for priming context to influence performance. By contrast, low-ability subjects have inefficient perceptual processes, so that orthographic input does not strongly activate semantic features for either high- or low-frequency targets. As a result, priming context influences the recognition of high- and low-frequency targets to a similar degree.

Planned t-tests were calculated to compare the related, unrelated, and nonword prime conditions in order to determine whether the priming effects resulted from facilitation, inhibition, or both (see Figure 2). Facilitation is indicated by faster recognition of a target following a related prime compared with a nonword prime, whereas inhibition is indicated by slower recognition of a target following an unrelated prime compared with a nonword prime (see Neely, 1991). The results indicated that there were no significant differences among priming conditions for the high-ability subjects reading high-frequency words, or for the low-ability subjects reading low-frequency words (ps $> .05$). However, the related prime condition was faster than both the unrelated and nonword prime conditions for the high-ability subjects reading low-frequency words, and the unrelated prime condition was slower than the nonword prime condition for the low-ability subjects reading high-frequency words (ps $< .05$). These effects suggest that priming resulted from facilitation as well as inhibition. The role of inhibition was further supported by the finding that RTs to target words were numerically slower after unrelated primes compared with nonword primes for all subjects regardless of their perceptual ability and for all targets regardless of their frequency. These findings support models of word recognition which predict that priming effects should result from facilitation and inhibition at long SOAs. As discussed in the Introduction, both the dual-mechanism and distributed network models predict this result, but they do so in very different ways (see Neely, 1977; Plaut, 1995; Stanovich, 1980).

The error analysis yielded a significant main effect of priming context, $F(1{,}468) = 8.46$, $MSE = 82.8$, $p < .001$, but this effect was qualified by an interaction between priming context and target frequency, $F(1{,}468) = 6.02$, $MSE = 59.0$, $p < .05$. The difference between the related and unrelated prime conditions was larger for low-frequency targets (1.0% vs 2.5%, respectively) than for high-frequency targets (1.2% vs 1.4%, respectively). This effect was produced primarily by facilitation through related primes and not by inhibition through unrelated primes, because the error rates for the unrelated prime conditions were less than 0.4% different from the error rates for the nonword prime conditions. Like the RT results reported above, this interaction can be accounted for in terms of the asymptotic effects depicted in Figure 1. In contrast to the RT results, the interaction between priming context and target frequency was not modulated by differences in perceptual ability, probably because the mean accuracy rates were above 96% for all ability levels.

In summary, when adult subjects performed lexical decision to target words following primes at a long SOA (800 ms), only the subjects with high perceptual ability—as assessed by a match-to-sample pretest—exhibited the standard finding of greater priming for low-frequency targets compared with high-frequency targets. Subjects with low perceptual ability exhibited equivalent levels of priming for low- and high-frequency targets. Moreover, by comparison with a nonword prime baseline, these effects were due to a combination of facilitation from related primes and inhibition from unrelated primes.

Experiment 2

The main purpose of Experiment 2 was to determine whether the priming differences observed among college students varying in perceptual ability in Experiment 1 could be replicated in a population of third and sixth grade children. We expected that elementary students with high perceptual ability would use priming context more for recognizing low-frequency words than for recognizing high-frequency words, whereas those with low perceptual ability would not exhibit an interaction between priming context and target frequency.

A second goal of Experiment 2 was to investigate whether priming context effects in elementary students result from facilitation, inhibition, or both. Experiment 1 revealed that the recognition of target words by college students at a long SOA was influenced both by facilitation from related primes and by inhibition from unrelated primes, relative to a nonword prime condition. This is consistent with past literature on semantic priming in adults at long SOAs (see Neely, 1991, for a review). Based on the Simpson and Lorsbach (1983) study discussed in the Introduction, we expected that the priming effects exhibited by elementary students at a long, 800ms

SOA would reflect mostly facilitation rather than inhibition. Specifically, we predicted that elementary students would exhibit faster RTs to target words following related primes compared with nonword primes, but they would show no difference between RTs to target words following unrelated primes compared with nonword primes.

Our prediction that children should exhibit smaller inhibitory effects than adults is in direct contrast to the dual-mechanism, interactive compensatory model (Stanovich, 1980), which holds that age differences in priming result from developmental changes in the relative use of automatic spreading activation and expectancy-based processes. Specifically, this model argues that children use expectancy-based processes more than adults and, therefore, children should exhibit more inhibition. A distributed network model, by contrast, predicts greater inhibition for adults compared with children because adults have more reading experience. Recall that inhibition in such models arises from hysteresis in moving from the stable (attractor) pattern of an unrelated prime to that of a target. Within a network, additional training causes primes to be processed more strongly and thus generate more stable semantic representations.[10] Consequently, with more training, the network should show greater inhibition in moving from the representation of the prime to that of the target. A caveat on this prediction, however, is that nonword primes may also be processed more strongly by adults than by children, thereby possibly diminishing the increase of inhibition with age.

Method

Subjects. Subjects were 44 third graders (*M* age = 8.9; *SD* = 0.5) and 46 sixth graders (*M* age = 11.8; *SD* = 0.5) from 2 private schools in the metropolitan Washington DC area. All subjects had English as a first language and reported that their vision was corrected to normal.

Apparatus. Same as Experiment 1.

Materials and Design. Due to limits on the amount of time children could miss classroom activities, the priming task used in Experiment 2 was shortened for use with the elementary students based on college student performance in Experiment 1. Specifically, items were eliminated if their correct mean RT from Experiment 1 fell greater than 2.0 *SD*s above the correct mean RT for the related prime (7 items), unrelated prime (9 items), or nonword prime (9 items) conditions. Items were also eliminated if the correct mean RT to a nonword target fell greater than 2.5 *SD*s above the correct mean RT for the word prime (13 items) or nonword prime (19 items) conditions. Items with slow RTs were eliminated because they might have been too difficult for the elementary students. Items were also eliminated if their correct mean RT in the related prime condition was more than 10 ms greater than their correct mean RT in the unrelated prime condition (see Appendix 1). Since we were interested in the effect of a reliable priming context on the recognition of words varying in target frequency, removing items that did not yield a reliable priming effect was justified (Borowsky & Besner, 1993).

The resulting critical stimuli for the priming task in Experiment 2 were 72 prime-target pairs in each of 3 conditions: unrelated prime and word target pairs, related prime and word target pairs, and nonword prime and word target pairs. There were also 24 word prime and nonword target pairs and 24 nonword prime and nonword target pairs. These 5 conditions totaled 120 test pairs. All other aspects of the priming task were exactly the same as in Experiment 1.

The strength of association between the related prime and target word in Experiment 2 (*M* = .48, *SD* = .18) was very similar to that in Experiment 1 (*M* = .47, *SD* = .17). Frequency was dichotomized into high frequency (*M* = 242.2; *SD* = 173.0) and low frequency (*M* = 37.9; *SD* = 25.6). The high- and low-frequency mean values in Experiment 2 were less than 10 per million greater than the same values in Experiment 1 (Kučera & Francis, 1967). The similarity of these values makes it possible to compare results across the experiments, despite the fact that a different set of items were used.

The elementary students were also administered the PPVT-R and the Symbol Search Test (see Experiment 1).

Procedure. The testing procedure was the same as in Experiment 1. All subjects were tested within their school building during regularly scheduled school hours.

Results and Discussion

Subjects were dichotomized into a high- or low-perceptual-ability group based on their Symbol Search Test speed-accuracy score, and into a third-grade group or sixth-grade group (see Table 2). The low-ability group scored significantly poorer on the perceptual ability measure than the high-ability group, $t(90) = 11.99$, $p < .001$, and the third graders scored significantly poorer on the perceptual ability measure than the sixth graders, $t(90) = 9.17$, $p < .001$. The PPVT-R was only marginally correlated with the perceptual measure partialed for age suggesting that these instruments were measuring two distinct underlying abilities ($r = -.19$, $p = .07$). This also means that ability differences in priming should be interpreted as resulting from perceptual ability and not from vocabulary knowledge.

For the semantic priming task, an item with a correct mean RT of greater than 2.5 *SD*s from the overall mean in the related, unrelated, or nonword prime condition for

[10] A further contributing factor may be that children's semantic representations are less stable because they are less differentiated than those of adults (see, e.g., Keil, 1979, 1987).

Table 2

Means (and Standard Deviations) for Age and for the Perceptual and Vocabulary Ability Measures for Experiment 2

	Age (in years)	Perceptual Speed-Accuracy	Vocabulary Raw Score
Sixth Graders (N = 46)	11.8 (0.5)	3.2 (0.5)	133.7 (12.1)
Third Graders (N = 44)	8.9 (0.5)	4.4 (0.8)	110.0 (12.2)
High Perceptual Ability (N = 45)	11.3 (1.3)	3.1 (0.4)	129.5 (15.1)
Low Perceptual Ability (N = 45)	9.4 (1.2)	4.5 (0.7)	114.8 (15.7)

Note. Perceptual Speed-Accuracy is mean time (in seconds) divided by number correct (out of 45) on the Symbol Search Test of the WISC-III (Wechsler, 1991). Vocabulary Raw Score is the mean raw score (out of 171) on the PPVT-R (Dunn & Dunn, 1981).

either the high- or low-ability subjects was eliminated (1 item or 1.3%). Correct mean RTs or error rates for each word were entered into a 2 (perceptual ability: high, low) $\times$ 2 (priming context: related, unrelated) ANCOVA with target frequency as a continuous regressor variable (see Experiment 1 for description of the ANCOVA design and inferential analysis strategy). Figure 3 shows the mean RT for each condition. The RT analysis yielded significant main effects of perceptual ability, $F(1,284) = 148.09$, $MSE = 1{,}422{,}020$, $p < .001$, priming context, $F(1,284) = 12.00$, $MSE = 115{,}210$, $p < .001$, and target frequency, $F(1,284) = 17.80$, $MSE = 170{,}879$, $p < .001$. However, these main effects were qualified by a significant interaction between priming context, target frequency, and perceptual ability, $F(1,284) = 3.54$, $MSE = 24{,}414$, $p < .05$.

The context-by-frequency interaction was then analyzed separately for the high- and low-ability subjects. For the high-ability subjects, there was a significant main effect of priming context, $F(1,142) = 4.35$, $MSE = 40{,}822$, $p < .05$, but only a trend towards an effect of target frequency, $F(1,142) = 2.98$, $MSE = 27{,}940$, $p = .087$. For the low-ability subjects, there were significant main effects of priming context, $F(1,142) = 7.87$, $MSE = 77{,}271$, $p < .01$, and target frequency, $F(1,142) = 17.74$, $MSE = 174{,}263$, $p < .001$. As predicted, there was a significant interaction between priming context and target frequency for the high-ability subjects, $F(1,142) = 4.31$, $MSE = 31{,}062$, $p < .05$, but not for the low-ability subjects ($F < 1$). Only high-ability subjects showed larger priming effects for low-frequency words (d = 52 ms) than for high-frequency words (d = 17 ms). Low-ability subjects showed similar priming effects for both high- and low-frequency words (ds = 47 and 46 ms, respectively). This replicates the interaction between perceptual ability, priming context, and target frequency found for the college students in Experiment 1. In addition, the mean difference between the related and unrelated prime conditions was about 40 ms for the elementary students in Experiment 2, whereas this mean difference was about 20 ms for the college students in Experiment 1. This finding is consistent with the extensive literature showing that younger children exhibit larger priming effects than older children and adults (e.g., Schwantes, 1985, 1991; Stanovich, 1980; West & Stanovich, 1978). We could not test this assertion directly with the data from the current studies, however, because the college and elementary students were not administered the same items in the priming task.

The related, unrelated, and nonword prime conditions were then compared with planned t-tests in order to determine whether the priming context effects resulted from facilitation, inhibition, or both (see Figure 3). There were no significant priming context differences for the high-ability subjects reading high-frequency words ($p >$ .05). However, the related prime condition was faster than the unrelated and nonword prime conditions for the high-ability subjects reading low-frequency words, and for the low-ability subjects reading both low- and high-frequency words (ps $< .05$). these results suggest that priming effects resulted from facilitation to related target words and not from inhibition to unrelated target words.[11] Consistent with this, Simpson and Lorsbach (1983, 1987) found no evidence for inhibition in second- through sixth-grade students at a low relatedness proportion (25%; cf. 20% in the current experiment). As mentioned above, the interactive compensatory model predicts that younger and low-ability readers should use inhibitory expectancy-based processes to a greater degree than older and high-ability readers (Stanovich, 1980). This prediction is disconfirmed by our findings and those of Simpson and Lorsbach. The findings are more compatible with a distributed network account that predicts weaker inhibition earlier

[11] It is possible that, as discussed in the Method section of Experiment 1, the greater likelihood of word versus nonword targets following word primes may have led to an overestimate of facilitation relative to inhibition in the current experiment. This would, however, have required the children to pick up on fairly subtle prime-target contingencies, and there is evidence that younger children are relatively insensitive to experimental factors, such as relatedness proportion, that do influence priming effects in older children and adults (Simpson & Lorsbach, 1983).

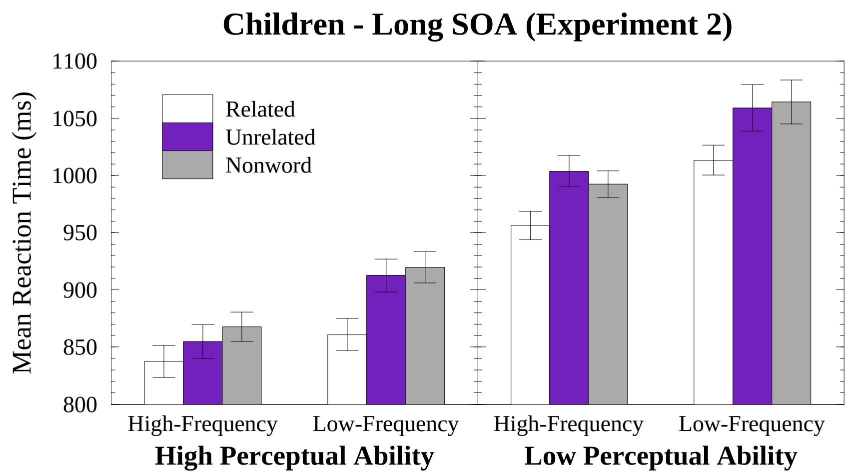

Figure 3. Item means for Experiment 2 of correct mean RTs to high- and low-frequency target words following related, unrelated, and nonword primes (800 ms SOA), for high- and low-perceptual-ability elementary students. Error bars indicate 1 standard error.

compared with later in training.

The error analysis yielded significant main effects of perceptual ability, $F(1,284) = 21.35$, $MSE = 522.4$, $p < .001$, priming context, $F(1,284) = 4.58$, $MSE = 112.1$, $p < .05$, and target frequency, $F(1,284) = 6.75$, $MSE = 165.3$, $p < .05$. High-ability subjects had lower error rates than low-ability subjects (1.8% vs 4.5%), a related priming context resulted in lower error rates than an unrelated priming context (2.5% vs 3.8%), and high-frequency targets had lower error rates than low-frequency targets (2.1% vs 4.3%). Whereas the college students in Experiment 1 exhibited a significant interaction between priming context and target frequency for error rates, the elementary students in the current experiment did not. This difference suggests that the frequency-by-context interaction for error rates becomes stronger with age, and supports the prediction that the interaction between priming context and target frequency should be stronger for older and high-ability readers than for younger and low-ability readers (see Figure 1). Again, though, differences in stimulus materials prevent a direct test of this claim.

So far we have only considered differences in priming due to perceptual ability. In order to examine age differences, correct mean RTs or error rates for each word were entered into a 2 (age: third-grade, sixth-grade) $\times$ 2 (priming context: related, unrelated) ANCOVA with target frequency as a continuous regressor variable. The RT analysis revealed significant main effects of age, $F(1,284) = 187.59$, $MSE = 1{,}805{,}640$, $p < .001$, priming context, $F(1,284) = 12.08$, $MSE = 116{,}267$, $p < .001$, and target frequency, $F(1,284) = 17.59$, $MSE = 169{,}347$, $p < .001$. The interaction between age, priming context, and target frequency was not significant, $F(1,284) = 1.86$, $MSE = 17{,}931$, $p = .17$. Because this three-way interaction was not significant, the interaction between priming context and target frequency was not analyzed separately for the third and sixth graders. The error analysis revealed significant main effects of age, $F(1,284) = 26.01$, $MSE = 677.5$, $p < .001$, priming context, $F(1,284) = 4.87$, $MSE = 126.9$, $p < .05$, and target frequency, $F(1,284) = 6.10$, $MSE = 158.9$, $p < .05$. Sixth grade subjects had lower error rates than third grade subjects (1.7% vs 4.8%, respectively; see the analysis of perceptual ability above for mean differences in priming context and target frequency effects)

In summary, like the adults in Experiment 1, when tested at a long SOA, the children in Experiment 2 exhibited greater priming for low- compared with high-frequency words only when they were of high perceptual ability. The low-ability children showed equal levels of priming regardless of target frequency. Moreover, by comparison with a nonword priming baseline, the children showed evidence of weaker inhibition than the adults, contrary to the predictions of the interactive compensatory model (Stanovich, 1980) but consistent with those of distributed network models. To our knowledge, no other developmental studies have used the single-word priming paradigm to determine whether the interaction between priming context and target frequency varies as a function of age; the few that have investigated this issue (e.g., Stanovich et al., 1981) have used a sentential priming paradigm. Experiment 2 also revealed a non-significant trend for an interaction between age, priming context, and target frequency. Age may not have explained a significant amount of variance in the interaction between priming context and target frequency because of individual differences in perceptual ability within the third and sixth graders (see Table 2). An important implication of this finding is that developmental researchers should reconsider using only age as an independent variable, at least in studies of word recognition. Since adults as well as children vary to a large degree in many different ways, one must consider individual differences in performance when interpreting developmental effects (see Epelboim, Booth, & Steinman, 1994, 1996).

Experiment 3

Considerable empirical work has focused on differences in the patterns of priming that result at short versus long SOAs. Effects that hold only at long SOAs are often ascribed to strategic, expectancy-based processes. Our first experiment established that perceptual ability in college students modulates the interaction between priming context and target frequency at a long, 800 ms SOA. The first goal of Experiment 3 was to determine whether this modulation also holds at a shorter, 200 ms SOA. As in Experiments 1 and 2, we predicted that, for high-ability subjects, priming context would have a greater influence on the recognition of low- compared with high-frequency targets, whereas, for low-ability subjects, priming context would have equal effects on low- and high-frequency targets.

The second goal of Experiment 3 was to determine whether the relative degree of facilitation and inhibition differs at short versus long SOAs. Although a number of previous studies have shown a predominance of facilitation at short SOAs but a combination of facilitation and inhibition at long SOAs (e.g., Becker, 1980; den Heyer et al., 1985b; Smith et al., 1987), none of these studies employed a nonword prime baseline. Our expectation, based on these earlier findings, is that priming effects should result primarily from facilitation at a short SOA. Dual-mechanism models (Neely, 1977, 1991) account for differential priming effects at short versus long SOAs in terms of separate mechanisms: Facilitation at a short SOA is due to automatic spreading activation, whereas the facilitation and inhibition at a long SOA are due to expectancy-based processes.[12] By contrast, we suggest that a distributed network model can account for the pattern of results based on the processing dynamics within a single mechanism. Recall that inhibition in such models results from hysteresis in moving from the stable pattern of the prime to that of the target. This inhibition should be weaker when the prime is processed less effectively and, thus, generates a less stable representation. Indeed, this was how we explained the weakened inhibition at a long SOA shown by the children in Experiment 2 compared with the adults at the same SOA in Experiment 1. By the same argument, inhibition should be weak or absent for adults at a short SOA, not due to the lack of an expectancy-based process, but because the prime has not had sufficient time to derive a stable representation (i.e., it hasn't settled very far into its attractor basin). In essence, on this account, we predict that adults at a short SOA should show similar priming effects to children at a long SOA.

[12] On some accounts, spreading activation also contributes to facilitation at long SOAs.

Method

Subjects. Subjects were 53 college students (*M* age = 19.1; *SD* = 2.1) at the University of Maryland who participated to fulfill a psychology course requirement. All subjects had English as a first language and reported that their vision was corrected to normal.

Apparatus. Same at Experiment 1.

Materials and Design. Same as Experiment 1, except that the SOA was 200 ms with 100 ms ISI.

Procedure. Same as Experiment 1.

Results and Discussion

Subjects were dichotomized into a high- or low-perceptual-ability group based on their Symbol Search Test speed-accuracy score (see Table 3). The low-perceptual-ability group scored significantly lower on the perceptual ability measure than the high-perceptual-ability group, $t(53) = 9.62$, $p < .001$. The vocabulary measure (PPVT-R) did not correlate significantly with the perceptual measure suggesting that these instruments were measuring two distinct underlying abilities ($r = .17$). This independence was supported further by the finding of no significant vocabulary ability differences between high and low perceptual ability groups ($t < 1$). These results indicate that any priming differences between the perceptual ability groups cannot be interpreted as being due to vocabulary differences.

For the semantic priming task, an item was eliminated from the analyses if it had a correct mean RT of greater than 2.5 *SD*s from the overall mean in the related, unrelated, or nonword prime condition for either the high- or low-perceptual-ability subjects (2 items or 1.6%). Correct mean RTs or error rates for each remaining word were entered into a 2 (perceptual ability: high, low) $\times$ 2 (priming context: related, unrelated) ANCOVA with target frequency as a continuous regressor variable (see Experiment 1 for a description of the ANCOVA design and inferential analysis strategy). The error analysis yielded no significant main effects nor interactions, so this analysis will not be reported. The mean RT for each condition is shown in Figure 4. The RT analysis yielded significant main effects of perceptual ability, $F(1,472) = 142.24$, $MSE = 763{,}699$, $p < .001$, priming context, $F(1,472) = 17.88$, $MSE = 96{,}005$, $p < .001$, and target frequency, $F(1,472) = 25.19$, $MSE = 135{,}262$, $p < .001$. However, these main effects were qualified by a significant three-way interac-

Table 3
Means (and Standard Deviations) for Age and for the Perceptual and Vocabulary Ability Measures for Experiment 3

	Age (in years)	Perceptual Speed-Accuracy	Vocabulary Raw Score
All Subjects (N = 53)	19.1 (2.1)	2.3 (0.3)	155.0 (12.4)
High Perceptual Ability (N = 26)	18.9 (2.1)	2.1 (0.1)	154.1 (13.8)
Low Perceptual Ability (N = 27)	19.3 (2.1)	2.6 (0.2)	156.0 (11.0)

Note. Perceptual Speed-Accuracy is mean time (in seconds) divided by number correct (out of 45) on the Symbol Search Test of the WISC-III (Wechsler, 1991). Vocabulary Raw Score is the mean raw score (out of 171) on the PPVT-R (Dunn & Dunn, 1981).

tion between perceptual ability, priming context, and target frequency, $F(1,472) = 4.35$, $MSE = 23{,}356$, $p < .05$.

The interaction between priming context and target frequency was then analyzed separately for the high- and low-ability subjects. For the high-ability subjects, there were significant main effects of priming context, $F(1,232) = 16.46$, $MSE = 67{,}143$, $p < .05$, and target frequency, $F(1,232) = 15.31$, $MSE = 62{,}423$, $p < .001$. For the low-ability subjects, there were also significant main effects of priming context, $F(1,240) = 4.81$, $MSE = 32{,}066$, $p < .05$, and target frequency, $F(1,240) = 10.97$, $MSE = 73{,}047$, $p < .01$. As predicted, there was an interaction between priming context and target frequency for the high-ability subjects, $F(1,232) = 9.99$, $MSE = 40{,}737$, $p < .01$, but not for the low-ability subjects ($F < 1$). Only high-ability subjects showed larger priming effects for low-frequency targets (d = 52 ms) than for high-frequency words (d = 17 ms). Low-ability subjects showed similar priming effects for high- and low-frequency targets (ds = 27 and 22 ms, respectively). In this way, Experiment 3, with a 200 ms SOA, replicated Experiments 1 and 2 with adults and children at long SOAs (800 ms).

Planned t-tests were then calculated to compare the related, unrelated, and nonword priming conditions in order to determine whether the priming effects resulted from facilitation, inhibition, or both (see Figure 4). The results indicated that the nonword priming condition was slower than the related and unrelated priming conditions for the high- and low-ability subjects reading high-frequency words (ps $< .05$). The related priming condition was faster than both the unrelated and nonword priming condition for the high-ability subjects reading low-frequency words, and the related prime condition was faster than the nonword prime condition for the low-ability subjects reading low-frequency words (ps $< .05$). Thus, the recognition of targets words was facilitated by related primes but not inhibited by unrelated primes.[13] These results, taken together with those of Experiment 1, support earlier empirical findings (Becker, 1980; den Heyer et al., 1985b; Neely, 1977, 1991; Smith et al., 1987) that only facilitation operates at short SOAs but both facilitation and inhibition operate at long SOAs.

Summary of Empirical Findings

The central question that motivated Experiments 1–3 is what underlies age and ability differences in the use of priming context for influencing visual word recognition (Schwantes, 1985, 1991; West & Stanovich, 1978). Our experiments revealed that subjects with high perceptual ability showed larger single-word priming effects for low-frequency targets than for high-frequency targets, but that subjects with low perceptual ability showed weaker and equal priming effects for high- and low- frequency targets. The interaction between priming context and target frequency was very consistent and robust. The high-ability subjects exhibited this interaction at short and long SOAs and whether they were elementary or college students.

We have explained the frequency-by-context interaction within the framework of distributed network models in terms of the asymptotic effect of the sigmoid activation function on the relative strength of orthographic input to semantics in the various conditions (Plaut, 1995, also see Cohen et al., 1990; Plaut et al., 1996, for analogous accounts of other interactive effects). Due to the increased training on high-frequency words, orthographic input for these words drives semantic units closer to the asymptote of the activation function, leaving a smaller range over which priming context can influence performance (see Figure 1). By contrast, low-frequency words produce input to semantics that falls closer to the linear range of the sigmoid where context effects are larger. Moreover, this framework can also explain why the frequency-by-context interaction holds only for subjects with high perceptual

[13] The current experiment may have overestimated facilitation relative to inhibition if subjects used prime-target contingencies to bias lexical decision (see the Methods section of Experiment 1). This is unlikely, however, given that subjects at short SOAs are generally insensitive to experimental manipulations that induce strategic effects (see Neely, 1991).

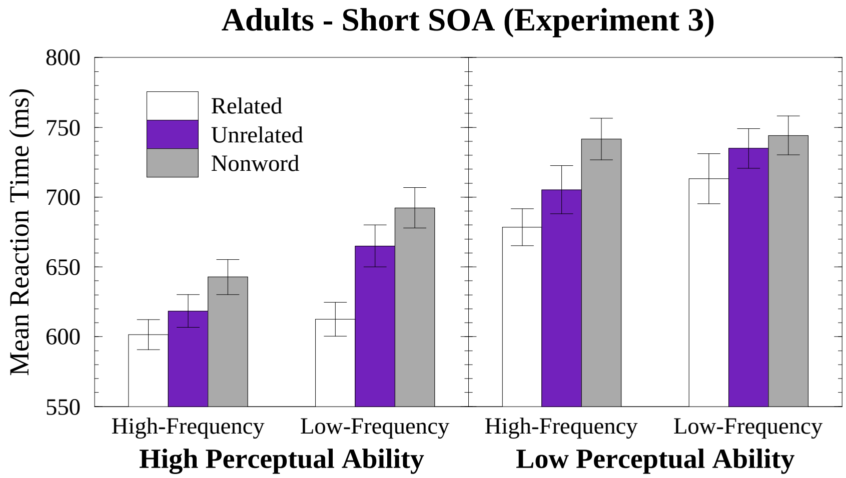

Figure 4. Item means for Experiment 3 of correct mean RTs to high- and low-frequency target words following related, unrelated, and nonword primes (200 ms SOA), for high- and low-perceptual-ability college students. Error bars indicate 1 standard error.

ability. If the orthographic input for low-ability subjects is insufficiently strong to drive even high-frequency targets into the nonlinear range of the sigmoid, both low- and high-frequency targets will fall within the linear range of the sigmoid where the effects of priming context are of equivalent magnitude.

A second purpose of the experiments was to determine the relative contribution of facilitation and inhibition to the priming effects at short and long SOAs when using a nonword prime baseline within a single-word priming paradigm. Experiments 1 and 3 showed that priming effects in adults resulted from only facilitation at a short SOA but from both facilitation and inhibition at a long SOA. These findings are consistent with those of other studies investigating facilitation and inhibition using single-word or sentence contexts with adults (e.g., Becker, 1980; den Heyer et al., 1985b; Smith et al., 1987; Stanovich & West, 1981). Experiment 2, by contrast, found that children show little if any inhibition at the same SOA at which adults show strong inhibition (also see Simpson & Lorsbach, 1983). If inhibition is equated with expectancy-based processes, the differences in inhibition between children and adults are inconsistent with the dual-mechanism interactive compensatory model (Stanovich, 1980) because it assumes that younger children depend more on expectancy-based processes than older children, and should therefore show more rather than less inhibition.

The distributed network framework, on the other hand, may provide a single-mechanism account of the pattern of facilitation and inhibition across SOAs shown by children and adults. Essentially, children at a long SOA do not exhibit inhibition for the same reason that adults at short SOAs do not—namely, the prime has been processed relatively weakly (due to weaker connection weights for the children, limited processing time for the adults) so that the network experiences little if any hysteresis in moving from the representation of the prime to that of the target. Expressed in terms of the concept of attractors, the network has not settled far into the attractor basin of the prime when the target is presented, so that there is little if any cost in moving up out of the prime's basin before settling to the attractor pattern for the target. Nonetheless, even in these conditions, a related prime still facilitates processing of the target more than an unrelated prime. This is because being even partway into the attractor basin for a related prime places the network in a state which has more overlap with the target (for categorical relatedness) and/or is closer to a state from which the network frequently moved to the target (for associative relatedness) as compared to an unrelated prime.

Computational Model

We claim that a distributed network model of semantic and associative priming can account both for the three-way interaction of priming context, target frequency, and perceptual ability, and for the patterns of facilitation and inhibition across age and SOA. If so, such a model would provide a more parsimonious account than dual-mechanism models (e.g,. Becker, 1980; Neely & Keefe, 1989; Stanovich, 1980) which must invoke distinct automatic, spreading activation processes and strategic, expectancy-based processes, and yet still have difficulty explaining the lack of inhibition for children at long SOAs. However, to this point, our claim is based solely on verbal characterizations of rather complex properties of distributed networks, only some of which have been demonstrated in existing simulations (e.g., the frequency-by-context interaction shown by Plaut, 1995).

To substantiate our account, and to demonstrate that a distributed network model can, in fact, exhibit the behavior we are ascribing to it, we developed a computational simulation of semantic and associative priming in lexical decision and applied it to the empirical findings from Experiments 1–3. The approach taken is closely related to the one used in the Plaut (1995) simulation.

Method

Stimuli. The actual stimuli used in the experiments were not employed in the simulation because of the complexity of their orthographic structure and because it was not feasible to derive realistic semantic representations for them. Rather, the network was trained on an abstract version of the task of mapping orthography to semantics. Although this task was simplified relative to the realistic mapping, it retained what we claim to be its essential characteristics.

Orthographic representations consisted of three-letter sequences constructed from 10 consonants (B, D, K, L, M, N, P, R, S, T) and 5 vowels (A, E, I, O, U). Letters were described in terms of six possible binary "features" such that each letter was assigned two of the six features. No attempt was made to assign similar codes to visually similar letters; codes were assigned to letters in alphabetic order. Words were restricted to consonant-vowel-consonant (CVC) strings; of the 500 possible CVC strings, 128 were chosen randomly to constitute the orthographic inputs on which the network was trained. Nonwords, by contrast, were restricted to VCV strings, with 128 chosen randomly out of the 250 possible such strings. Nonwords were constructed to be as different from the words as possible, in part because the word and nonword stimuli in the empirical studies were not orthographically matched (see Experiment 1 methods) and included many orthographically unwordlike nonwords (see Appendix 1). Note, however, that since vowel and consonant letters share features, there is some overlap between the orthographic representations for words and nonwords—just less than the overlap among the words themselves. Also note that the word and nonword stimuli are not necessarily words and nonwords in English. Rather, for the simulation, "words" are stimuli that are trained to activate a particular pattern of semantic features, and "nonwords" are novel stimuli that are never encountered during training but can nonetheless be presented to the network during testing.

The semantic representations of words were the same as those used by Plaut (1995). They were generated to cluster into artificial semantic "categories."[14] Eight different random binary patterns were generated over 100 semantic features, in which each feature had a probability p_a = .1 of being active. These patterns served as the prototypes for eight separate semantic categories. Sixteen exemplars were generated from each prototype pattern by randomly altering some of its features (Chauvin, 1988). Eight of these were high-dominance exemplars in which relatively few features of the prototype were changed (each feature had a probability of .2 of being resampled with p_a = .1). The remaining eight were low-dominance exemplars in which more features were altered (resampling probability of .4). In addition, all pairs of patterns were constrained to differ by at least four features. The effect of this manipulation is to make all exemplars within a category cluster around the prototype, with high-dominance exemplars more similar to the category prototype than low-dominance exemplars, and for all semantic patterns to have an average of 10 active features (range 4–18) out of a total of 100. Although the effect of target category dominance was not explored in the current work, the Plaut (1995) simulation exhibited greater semantic priming of high- compared with low-dominance targets, in keeping with empirical findings (Lorch, Balota, & Stamm, 1986; Schwanenflugel & Rey, 1986).

Semantic patterns were assigned to orthographic patterns randomly to ensure, as is true of monomorphemic words in English, that there was no systematic relationship between orthography and semantics. Words were considered semantically related if their semantics were generated from the same prototype. Half of the words in each category were designated as high-frequency and, as described below, were presented twice as often during training as the remaining, low-frequency words.

Network Architecture. The architecture of the network is shown in Figure 5. Eighteen *orthographic* units (three banks of six features) encoded the three-letter input. These units were fully connected to 100 *hidden* units which, in turn, were fully connected to 100 *semantic* units. The semantic units were fully connected to each other as well as back to the hidden units. In addition, each hidden and semantic unit had a *bias* connection from a unit whose activity was always 1.0; the weight on this connection can be thought of as determining the unit's "resting" activation in the absence of other input, and is learned in the same way as the other weights in the network. Including biases, the network had a total of 34,024 connections. The weights on connections were initialized to random values sampled uniformly between ± 0.25.

The states of units in the network change smoothly and continuously in time in response to influences from other units. For the purposes of simulation on a digi-

[14]We characterize the basis for similarity (i.e., feature overlap) among semantic representations in terms of categories for ease of exposition. Note, though, that the general concept of semantic relatedness extends beyond simple taxonomic category membership (see, e.g., Moss, Ostrin, Tyler, & Marslen-Wilson, 1995).

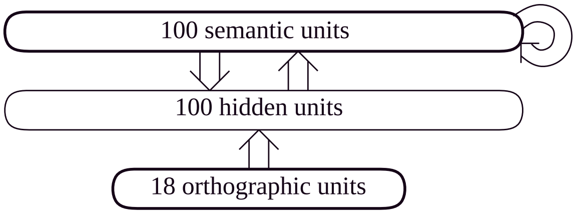

Figure 5. The architecture of the network. Ovals represent groups of units and arrows represent full connectivity between these groups.

tal computer, it is convenient to approximate continuous units with finite difference equations in which time is discretized into *ticks* of some duration τ. Thus, the activation of unit j at time t is given by

$$a_j^{[t]} = \tau\, \sigma\!\left(\sum_i w_{ij} a_i^{[t-\tau]} \right) + (1-\tau)\, a_j^{[t-\tau]} \qquad (1)$$

where w_{ij} is the weight from unit i to unit j and $\sigma(x) = (1 + \exp(-x))^{-1}$ is the standard sigmoid function (the top portion of which is depicted in Figure 1). According to this equation, a unit's activation at each point in time is a weighted average of its current activation and the one dictated by other units, where τ is the weighting proportion. As τ approaches zero, the discrete system more closely approximates the true underlying continuous system but the demands on computational resources increase. In the current simulation, a relatively large value of τ was used during training, when minimizing computation time is critical, whereas a much smaller value of τ was used during testing, when a more precise measure of settling time was required. Note, though, that this manipulation of τ did not significantly change the settling behavior of the network nor the final pattern it produced for any given input.

Training Procedure. The network was trained in the following way. On most training trials, units started with the activations they had at the end of processing the previous word. However, for the very first word, and with a probability of .01 throughout training, the activations of units were initialized to 0.2. This was done to insure that the network was capable of processing words correctly in the absence of context from a preceding word, as was the case for the presentation of primes in the testing procedure. A word was presented to the network by providing each orthographic unit with external input which was positive if the corresponding orthographic feature was present in the word's representation and negative if it was absent. The strength of this external input—controlled by a parameter termed *input strength*—in intended to reflect the relative effectiveness of lower-level perceptual processes not implemented in the current simulation. Specifically, the input strength parameter specifies the fraction of the distance from the neutral activation of 0.2 to the relevant extreme of the sigmoid function (1.0 for present features and 0.0 for absent features) that would be produced by the external input in the absence of other input to a unit.[15] This external input remained constant during the processing of the word. Although the input strength parameter was held constant at 0.8 during training, it was manipulated during the testing of the network to model the performance of subjects with high versus low perceptual ability, as described below.

Given the external input resulting from the presentation of a word, all of the units in the network (including the orthographic units) updated their states according to Equation 1 for every time *tick* of duration τ = 0.2 over a total of 4.0 units of time. (Note that the absolute time scale of the network is determined by the time constants of the underlying differential equations, which were assumed to be equal to 1.0.) The performance error of the network was measured by the *cross-entropy*, C (Hinton, 1989), between the activations of the semantic units, a_j, and the assigned semantic pattern for the presented word, s_j, throughout the last unit of time.

$$C = \tau \sum_{3<t\leq 4} \sum_j s_j \log\!\left(a_j^{[t]}\right) + (1-s_j) \log\!\left(1 - a_j^{[t]}\right) \qquad (2)$$

A continuous version of back-propagation through time (Pearlmutter, 1989) was then used to calculate the partial derivative of this measure with respect to each weight in the network. The weights were updated immediately after each word presentation i according to

$$\Delta w_{ij}(i) = \epsilon \frac{\partial C}{\partial w_{ij}} + \alpha \Delta w_{ij}(i-1) \qquad (3)$$

using a learning rate ϵ = 0.005 and momentum α = 0.8. Although the network received error only during the last of four units of time, the back-propagated error exerted a pressure on the network to settle to the correct semantic pattern as quickly as possible.

The selection of the next word for training was determined both by relatedness and by frequency. Given the high co-occurrence of semantic and associative relatedness in the experimental stimuli (see Appendix 1) and in natural language (see, e.g., Postman & Keppel, 1970), all word pairs that were semantically related (i.e., in the same category) were also made associatively related by increasing their transition probabilities (Plaut, 1995).[16] Specifically, with a probability of 1/7, the next word to be

[15] For example, an input strength of 0.8—the value used during training—specifies an external input of 0.575 for present features and -3.18 for absent features, because $0.8 \times (1.0 - 0.2) = 0.64 = \sigma(0.575)$ and $0.2 - 0.8 \times (0.2 - 0.0) = 0.04 = \sigma(-3.18)$.

[16] The complete consistency between semantic and associative relatedness employed in the current simulation is certainly an exaggeration

presented was selected randomly from among the other words in the same semantic category as the current word; on the remaining 6/7 of trials, the next word was selected randomly from the entire set of words. The value of 1/7 was chosen so that the next word was twice as likely to come from the same category as from another category. In addition, high-frequency words were twice as likely to be trained as low-frequency words. In this way, the relative influence of frequency and semantic/associative relatedness were equated in the network.

The network was trained for a total of 200,000 word presentations, at which point it was completely accurate in settling into the semantic representation of each word regardless of the preceding context. However, as described below, we also examined the performance of the network at an earlier point in training (70,000 word presentations) as an approximation to the reading experience received by the children in the empirical studies.

Testing Procedure. During testing, stimuli were presented to the network in prime-target pairs. First, the network was initialized to activations of 0.2. Then the prime was presented, as external input to the orthographic units, and processed for some specified duration. The prime was then replaced by a "blank" input (all zeros) and processing continued until some specified stimulus-onset asynchrony (SOA) had elapsed. Following this, the target was presented and the network continued processing until the semantic activation stopped changing—specifically, until the activation of each semantic unit differed from the sigmoid of its summed input from other units (see Equation 1) by no more than 0.05. At this point, the network was considered to have responded and the time elapsed since the presentation of the target was taken as its reaction time (RT). In order to compute these RTs precisely, the network was tested using a much finer temporal discretization (τ = 0.01) than was used during training ($\tau = 0.2$). As mentioned earlier, however, this manipulation had a negligible effect on the final activations produced by the network.

The network was run under eight conditions by fully crossing three factors. The factor *age* specified the amount of training experienced by the network—either 70,000 word presentations (child) or 200,000 word presentations (adult). The value of 70,000 was chosen so that the relative difference in overall RT between the child and adult conditions was approximately the same for the network as for the human subjects. The second factor, *perceptual ability*, was instantiated by increasing the input strength parameter from the value of 0.8 used during training—either slightly, to 0.82 (low ability) or more extensively, to 0.9 (high ability). These values were chosen to approximate the relative difference in performance between the high- and low-ability subjects in the empirical studies. The third factor, *SOA*, reflected the timing of stimuli—prime-target pairs were presented either at an SOA of 1.0 unit of time with an ISI of 0.5 (short SOA) or at an SOA of 4.0 with an ISI of 1.0 (long SOA). Note that the SOA and ISI values are directly proportional to those used in the empirical studies (short SOA of 200 ms with 100 ms ISI; long SOA of 800 ms with 200 ms ISI).

For each level of age, perceptual ability, and SOA, the RTs for each word and nonword as target were measured when preceded by every other word and nonword as prime.[17] Lexical decisions were based on a measure of the familiarity of the resulting semantic pattern (Atkinson & Juola, 1973; Balota & Chumbley, 1984). The specific measure of familiarity that was used is termed *semantic stress* and reflects the degree to which semantic activations are binary (also see Plaut, 1997). More formally, the stress S_j of unit j is a measure of the information content (entropy) of its activation a_j, corresponding to the degree to which it differs from the "neutral" output of 0.5 (the value generated by the sigmoid given zero input):

$$S_j = a_j \log_2(a_j) + (1 - a_j)\log_2(1 - a_j) - \log_2(0.5) \quad (4)$$

The stress of a unit is 0 when its activation equals 0.5 and approaches 1 as its activation approaches either 0 or 1. Because, over the course of training, the semantic patterns generated by words come to approximate their binary target patterns, the average semantic stress for words approaches 1. By contrast, nonwords are novel stimuli that typically fail to drive semantic units as strongly as words do, resulting in much lower semantic stress values. We assume that subjects can adopt a decision criterion that optimally distinguishes words from nonwords on the basis of the distribution of semantic stress values. Plaut (1997) showed that, under this assumption, semantic stress provided a reliable basis for lexical decision in a feedforward network that was trained to map from orthography to semantics for the 2998 words in the Plaut et al. (1996) training corpus. For the network, presentations of word targets generating stress values below the decision criterion were considered errors and were not included in the RT analyses.

For each word target in each network condition, three item means were computed: 1) the mean RT of correct responses to the target word when preceded by each of the 15 non-identical primes which were both associatively

of the actual co-occurrence of these factors. Even so, because there was no attempt to dissociate these factors in the current empirical work, this approximation was considered adequate in the current context.

[17] Note that the network was reinitialized before each prime presentation, so there is no possibility of cross-trial contaminating effects due to target or prime repetition, as there would be with human subjects. It should also be pointed out that this testing procedure involves 524,280 target presentations and, assuming an average RT of 5.0 at τ = 0.01, requires well over a week of CPU time on a processor running at 30 Mflops.

and categorically related to it (*related* condition); 2) the mean RT of correct responses to the target word when preceded by each of the 112 primes which were neither associatively nor categorically related to it (*unrelated* condition); and 3) the mean RT of correct responses to the target word when preceded by each of the 128 nonword primes (*nonword* condition).

In summary, the network was tested in a fully crossed, five-factor design involving age (child vs. adult), perceptual ability (high vs. low), SOA (short vs. long), target frequency (high vs. low), and priming context (related vs. unrelated vs. nonword), for a total of 48 cells. Target frequency is a between-item factor whereas all others are within-item factors.

Results and Discussion

Figure 6 shows the distribution of semantic stress values produced by word and nonword targets in the child and adult conditions, averaged over all other conditions and prime times. When compared with the child condition, the adult condition produces higher stress values, particularly for word targets, and much less overlap between the word and nonword distributions. Given the strong effect of age (i.e., amount of training) on the discriminability of words and nonwords, separate decision criteria were applied in performing lexical decision in the adult and child conditions. These criteria, shown as vertical lines in the figure, were chosen to roughly approximate the proportion of hits to false-alarms exhibited by the corresponding subjects in the empirical studies. Specifically, for the adult condition, making a “yes” response when semantic stress equals or exceeds a criterion of 0.91 yields 99.8% hits and 0.32% false-alarms; the corresponding values were 98% and 6.8% for the adult subjects in Experiments 1 and 3. For the child condition, a criterion of 0.87 yields 98.3% hits and 10.3% false-alarms (cf. 97% and 11% for the children in Experiment 2). Thus, overall, the network is slightly more accurate at lexical decision than are the subjects, particularly in the adult condition.

For all correct responses to word targets (hits), RTs outside ± 2.5 *SD*s within each cell of the design (i.e., each combination of age, perceptual ability, SOA, target frequency, and priming context) were withheld from the latency analysis. Overall, this removed 2.8% of the observations, with a maximum over cells of 5.4%. In general, there was slightly greater trimming for cells involving low perceptual ability and low target frequency.

To facilitate comparison of the network’s performance with that of the subjects’, the network’s RTs were converted to milliseconds by computing a mean RT for each condition and then linearly regressing these means against the corresponding 36 condition means for the subjects’ from Experiments 1–3. Note that the experiments did not involve testing children at a short SOA and, therefore, the

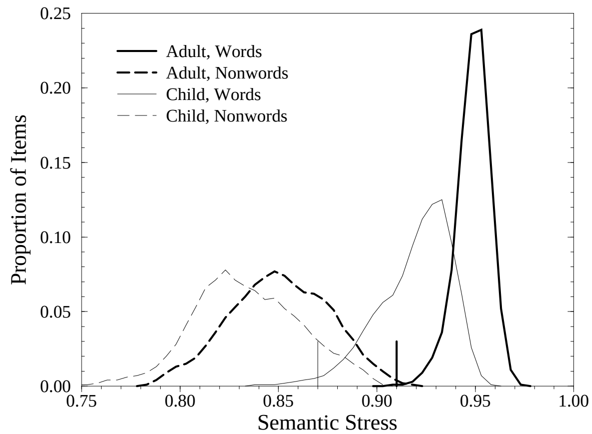

Figure 6. Distribution of semantic stress values produced by the network for word targets (solid lines) and nonword targets (dashed lines) in the child condition (70,000 word presentations; light lines) and in the adult condition (200,000 word presentations; dark lines). The small vertical lines indicate the decision criteria used for lexical decision for the child conditions (left, light line) and the adult conditions (right, dark line).

mean RTs of the network under these conditions could not be entered into the regression equation. Nonetheless, these mean RTs were scaled to milliseconds using the regression equation for the 36 condition means. The resulting values constitute empirical predictions of subject mean RTs for children at the short SOA (see below). Also note that, by applying a single regression equation to convert network RTs to milliseconds, we are implicitly testing the very strong hypothesis that a single network, under specific manipulations for age and perceptual ability, can account for the effects of target frequency, priming context, and SOA across six different groups of subjects (i.e., 2 levels of perceptual ability in each of 3 experiments).

Based on the linear regression, each network RT, t, was converted to milliseconds by the equation $325.2\,t - 774.825$. Figure 7 displays the condition means for both the subjects and the network; the numeric values of the means and their standard deviations are given in Appendix 2. The overall correlation of the network and subject mean RTs across the 36 conditions was high: r = .92 (r^2 = .85). As the figure shows, the main points of discrepancy concern the child conditions. In particular, the network over-estimates the frequency effect for both high- and low-ability children at the long SOA. The same tendency is also apparent to some extent for the low-ability adults at the long SOA. The relatively stronger effect of frequency for the network compared with the subjects may be due in part to the fact that frequency was a strongly dichotomous variable for the network—high-frequency words were always twice as frequent as low-frequency words—whereas frequency was treated as a continuous variable in the empirical analyses. Moreover, the empirical study with children used only moderately low-frequency words to ensure a high level of accuracy (see Experiment 2 methods). With a stronger frequency manipulation than that used in Experiment 2, Perfetti and Hogaboam (1975) found that, among children, the poor readers exhibited larger frequency effects than did the good readers.

Given the conversion of network RTs to milliseconds, we carried out analyses on the subsets of the network's data that corresponded to the conditions in each of the three experiments. We then performed additional analyses on aspects of the network's performance that constitute empirical predictions not directly tested by the experiments. In these analyses, frequency was treated as a dichotomous variable because it was instantiated as such in the simulation. For any interactions with perceptual ability, planned comparisons were used to determine if the relevant effects were reliable within each ability group. In addition, planned comparisons between the related, unrelated, and nonword conditions were used to determine if the priming effects were due to facilitation, inhibition, or both. All of the analyses are over items (N = 128) and, therefore, all F-tests have degrees of freedom of (1,126). Lastly, it was noticed that the standard errors for the estimates of the network's condition means were generally much smaller than those of the subjects, particularly for the conditions involving high perceptual ability. This is presumably because item means for the network were calculated over a larger number of equivalent primes, and because processing in the network was not subject to any intrinsic variability. Therefore, a more stringent reliability criterion was adopted for the network; specifically, statistical tests were considered reliable only at $p < .01$.

The use of fixed criteria for determining settling times and lexical decisions meant that the network could not vary its speed-accuracy tradeoff across conditions (although see the General Discussion). Moreover, given the high discriminability of words and nonwords, particularly for the adult conditions (see Figure 6), many testing conditions produced no errors. Consequently, there were relatively few reliable effects in the analysis of errors produced by the network, and those effects which did hold were all in the same direction as those found in the RT analyses reported below. In fact, across the 36 experimental conditions, there was a high correlation (r = .74, $p < .001$) between mean RT and error rate. Moreover, the only reliable interaction in the error data from the empirical studies was between target frequency and priming context for adults at the long SOA (Experiment 1), and this was in the same direction as the effect in RTs. Given these considerations, only RT analyses for the network are presented below.

Simulation of Experiment 1: Adults, Long SOA. Analogous to Experiment 1, correct mean RTs of the network for each word in the adult condition (200,000 word presentations) tested at the long SOA (SOA = 4.0, ISI = 1.0) were entered into a three-factor ANOVA with perceptual ability (high vs. low) and priming context (related vs. unrelated) as within-item factors and target frequency (high vs. low) as a between-item factor. Figure 8 shows the means for each condition both for the subjects in Experiment 1 and for the network. The analysis revealed clear main effects of perceptual ability (F = 61.73, MSE = 7239, $p < .001$) target frequency (F = 47.33, MSE = 11,766, $p < .001$) and priming context (F = 75.62, MSE = 1746, $p < .001$), as well as a reliable interaction between ability and frequency (F = 9.59, MSE = 7240, $p < .005$). However, these effects were qualified by a reliable three-way interaction of ability, frequency, and context (F = 10.26, MSE = 1262, $p < .005$). These findings agree with those of Experiment 1, except for the two-way interaction of ability and frequency. In the network, the frequency effect is smaller in the high- versus low-ability condition (ds = 43 and 89 ms, respectively). For the subjects, this difference (ds = 31 and 40 ms) was in the same direction numerically but was not reliable.

Separate analyses for the high- and low-ability conditions in the network revealed reliable frequency and context effects for both (Fs $\geq$ 29.64, ps $< .001$). As in the empirical findings of Experiment 1, there was a reliable frequency-by-context interaction in the high-ability condition (F = 16.98, MSE = 230.1, $p < .001$) but not in the low-ability condition (F = 3.49, MSE = 2778, p = .064). Interestingly, as can be seen in Figure 8, there was a trend in the low-ability condition in the opposite direction to the interaction for the high-ability condition. Specifically, in the high-ability condition, the network exhibited less priming for high-frequency targets than for low-frequency targets (ds = 15 and 29 ms, respectively) which agrees with the empirical findings (ds = 7 and 33 ms). In the low-ability condition, the network's priming effect was larger for high- versus low-frequency targets (ds = 54 and 28 ms, respectively). The corresponding numeric difference for subjects (ds = 20 and 17 ms) was in the same direction but was not reliable.

We have explained the standard finding of greater priming for low-frequency targets in the high-ability condition in terms of the nonlinearity of sigmoid activation function (see Figure 1). In fact, the same principles can explain a trend towards the reverse interaction—greater priming for high-frequency targets—in the low-ability condition. If the bottom-up contribution of perceptual ability to the input of semantic units is sufficiently weak, the low-frequency targets may start to fall within the opposite tail of the sigmoid—this is, in fact, reflected in Figure 1. As a result, the effects of priming context would be reduced

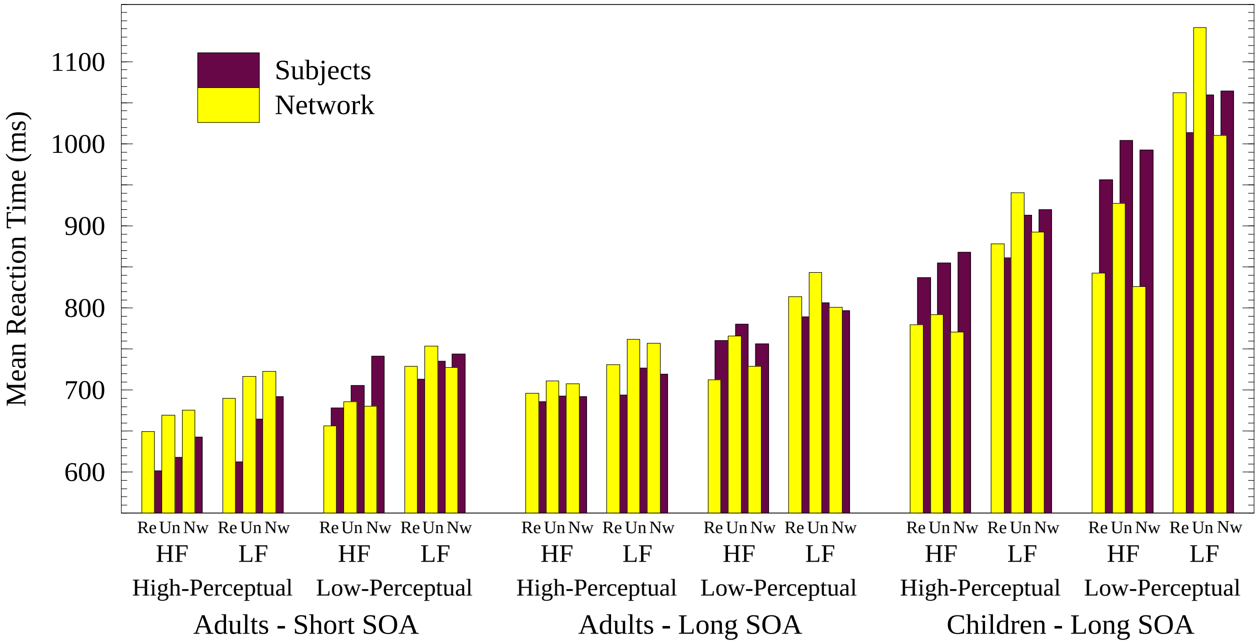

Figure 7. The mean RTs for subjects from Experiments 1–3 and the corresponding mean RTs for the network, converted to milliseconds by regressing the network's mean settling times across all 36 conditions against the subject means. "Re" is for related primes, "Un" is for unrelated primes, "Nw" is for nonword primes, "HF" is for high-frequency targets, and "LF" is for low-frequency targets.

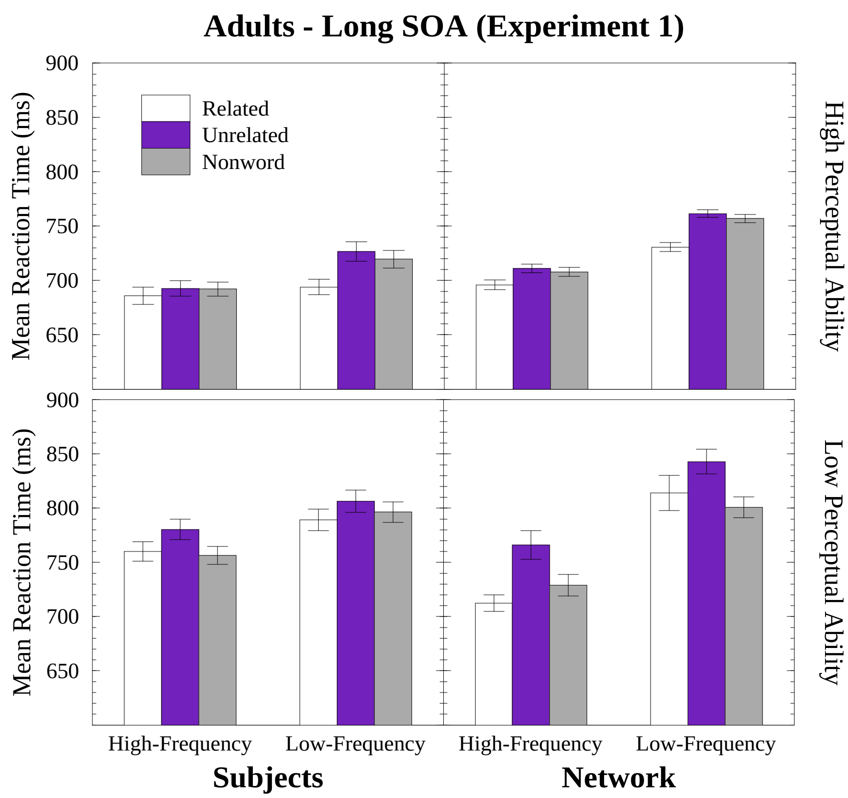

Figure 8. Mean RTs for high- and low-perceptual-ability adult subjects at the long SOA (from Experiment 1) and the mean RTs for the network from the corresponding conditions. Error bars are 1 standard error by items.

relative to those for high-frequency targets which still fall within the linear range of the function.[18] However, given that this reverse frequency-by-context interaction was not reliable in either the empirical nor simulation data, this account should be considered merely suggestive until the finding is verified by additional empirical and computational investigation.

As in the empirical analyses, planned comparisons with the nonword prime condition within each combination of target frequency and perceptual ability were calculated to determine whether context effects were due to facilitation (related vs. nonword conditions) and/or inhibition (unrelated vs. nonword conditions). In the high-ability conditions, context effects were due mostly to facilitation (*p*s $< .001$) with only marginal inhibition ($p = .015$ for low-frequency targets; $p = .078$ for high-frequency targets). In the low-ability conditions, context effects were due to both facilitation ($p < .014$) and inhibition ($p < .001$) for high-frequency targets, but only inhibition for low-frequency targets ($p < .001$). These findings are broadly consistent with those from Experiment 1, except for the findings of reliable facilitation for high-frequency targets in the high-ability condition and reliable inhibition for low-frequency targets in the low-ability condition. In both of these cases, however, the numeric differences in the empirical data agree with the effects in the network. Also, as was true of the empirical findings, mean RTs for unrelated primes were numerically slower than for nonword primes regardless of target frequency and perceptual ability.

In summary, the pattern of RTs produced by the network in the adult, long-SOA conditions matches fairly closely the pattern of results produced in Experiment 1 by adults tested at the long (800 ms) SOA. Most important, the network exhibited the appropriate three-way interaction of perceptual ability, target frequency, and priming context, with greater priming for low-frequency targets only in the high-ability condition. In fact, the low-ability condition showed a trend towards greater priming for high-frequency targets. While this reverse interaction held numerically in the subject data and can be understood within the general framework of distributed network models, it is in need of further verification. Finally, the network also produced the observed empirical pattern of a mixture of facilitation and inhibition at the long SOA, with primarily facilitation in the high-ability conditions and primarily inhibition in the low-ability conditions.

Simulation of Experiment 2: Children, Long SOA. An ANOVA analogous to the one just described was carried out on the network's correct mean RTs for each word in the conditions corresponding to Experiment 2—the child, long-SOA conditions (70,000 word presentations; SOA = 4.0, ISI = 1.0). Figure 9 shows the means for each condition both for the subjects in Experiment 2 and for the network. Although, like the subject data, the network's RTs at the long SOA were much slower in the child condition (mean 920 ms) than in the adult condition (mean 754 ms), the pattern of results was quite similar. There were reliable main effects of perceptual ability ($F = 159.60$, $MSE = 17{,}051$, $p < .001$), target frequency ($F = 68.83$, $MSE = 53{,}190$, $p < .001$), and priming context ($F = 125.69$, $MSE = 3626$, $p < .001$). Perceptual ability also interacted with target frequency ($F = 16.30$, $MSE = 17{,}051$, $p < .001$), and with priming context ($F = 23.70$, $MSE = 2698$, $p < .001$). These interactions were not reliable in the empirical data from Experiment 2 but both were in the same direction numerically as in the network's data. The network's two-way interactions were, however, qualified by a three-way interaction of ability, frequency, and context ($F = 8.89$, $MSE = 2698$, $p < .005$).

There were reliable frequency and context effects in both the high- and low-ability conditions when analyzed separately (*F*s ≥ 53.27, *p*s $< .001$). Moreover, as in the corresponding analysis of the empirical data from Experiment 2, the frequency-by-context interaction was reliable in the high-ability condition ($F = 29.24$, $MSE = 1345$, $p < .001$) but not in the low-ability condition, ($F < 1$).

The relative contributions of facilitation and inhibition to context effects were determined by planned comparisons of the mean RTs for the nonword prime conditions to the related and unrelated conditions. In the high-ability condition, there was only facilitation for high-frequency targets ($p < .001$) but both facilitation and inhibition for low-frequency targets (*p*s $< .001$). In the low-ability condition, there was both facilitation and inhibition for high-frequency targets (*p*s $< .01$) but only inhibition for low-frequency targets ($p < .001$). As Figure 9 makes clear, the main discrepancies between the network findings and those in Experiment 2 are due to the nonword priming conditions for the low-ability subjects. For both high- and low-frequency targets, the network exhibited clear inhibition—faster responses following nonword primes as compared with unrelated primes—whereas the subjects

[18] It should be pointed out, though, that a direct translation from the output values of the sigmoid activation function to the settling times in a fully recurrent distributed network is a considerable simplification, particularly when relying on the opposite side of the sigmoid. activation of a unit. There is, however, an alternative explanation for a reverse frequency-by-context interaction in the low-ability condition that relies on the same side of the sigmoid as the standard interaction. Given that the network is exposed to high-frequency items more often than to low-frequency items during training, it follows that the bottom-up contribution of context effects, particularly for associative relatedness, should be stronger for high- compared with low-frequency targets. As a result, if both high- and low-frequency targets fell within the linear range of the sigmoid function, the high-frequency targets would exhibit greater priming effects. This reverse interaction would be eliminated, however—as it is in the current empirical and computational data—if the input contribution in the low-ability condition was slightly stronger, so that high-frequency targets fell slightly into the side of the sigmoid that leads to the standard interaction in the high-ability condition.

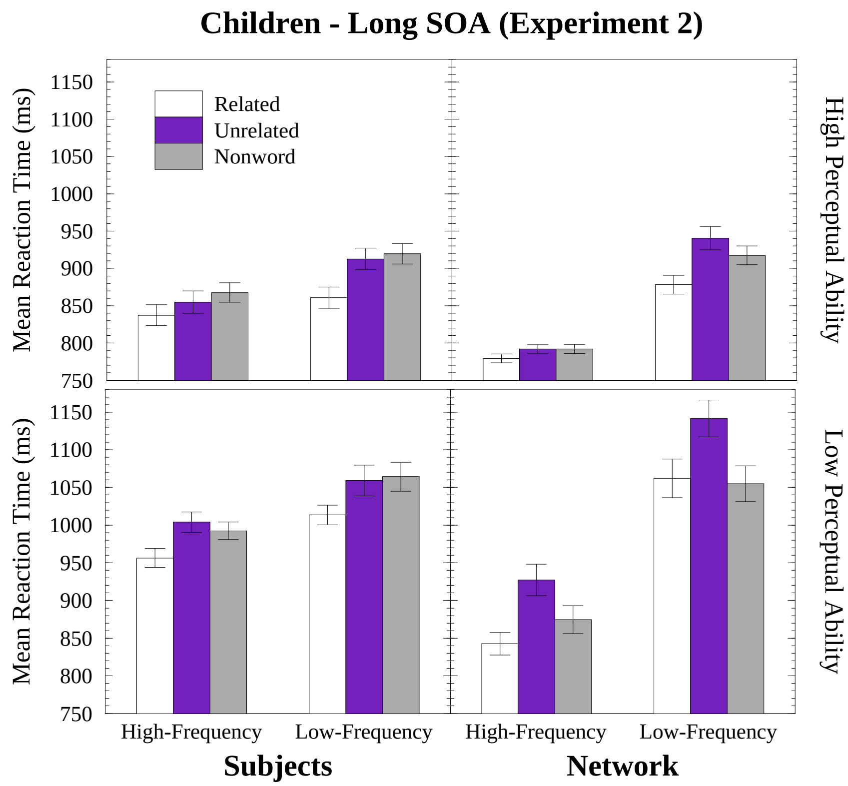

Figure 9. Mean RTs of high- and low-perceptual-ability children at the long SOA (from Experiment 2) and the mean RTs for the network from the corresponding conditions. Error bars are 1 standard error by items.

showed none. In fact, even low-frequency targets in the high-ability condition produced a trend towards inhibition whereas the subjects did not. We will consider the implications of these discrepancies in the General Discussion.

In summary, as in the empirical studies, the network in the child, long-SOA conditions produced a pattern of priming rather similar to that of the adult condition at the long SOA. Most important, the standard finding of greater priming for low-frequency targets held only in the high-ability condition; high- and low-frequency targets produced nearly equal levels of priming in the low-ability condition. The relative contribution of facilitation and inhibition was, however, somewhat different for the network than for the subjects, particularly for the conditions involving low perceptual ability.

Simulation of Experiment 3: Adults, Short SOA. Analogous to Experiment 3, another ANOVA was carried out on the correct mean RTs of the network for each word in the adult, short-SOA conditions (200,000 word presentations; SOA = 1.0, ISI = 0.5). Figure 10 shows the means for each condition both for the subjects in Experiment 3 and for the network. Consistent with the empirical findings, the network analysis showed main effects of perceptual ability (F = 20.79, MSE = 3825, $p < .001$), target frequency (F = 54.10, MSE = 7697, $p < .001$), and priming context (F = 121.37, MSE = 670.5, $p < .001$). These main effects were qualified, however, by a three-way interaction of ability, frequency, and context (F = 7.52, MSE = 165.1, $p < .01$).

There were reliable effects of target frequency and priming context in both the high- and low-ability conditions when analyzed separately (Fs $\geq$ 33.85, ps $< .001$). The frequency-by-context interaction was marginal in the high-ability condition (F = 3.39 MSE = 259.7, p = .068) but was not reliable in the low-ability condition ($F < 1$). Thus, as in the corresponding empirical data, the network tended to produce greater priming for low- compared with high-frequency targets in the high-ability condition (ds = 27 and 19 ms, respectively), but equivalent amounts of priming for low- and high-frequency targets in the low-ability condition (ds = 25 and 30 ms). The fact that, in the low-ability condition, priming for low-frequency targets was numerically smaller than for high-frequency targets agrees with the empirical findings from Experiment 3 (ds = 22 and 27 ms) and echoes the trend towards a reverse frequency-by-context interaction discussed in the context of the simulation of adults at the long SOA (Experiment 1).

Planned comparisons of the related and unrelated prime conditions with the nonword prime condition within each

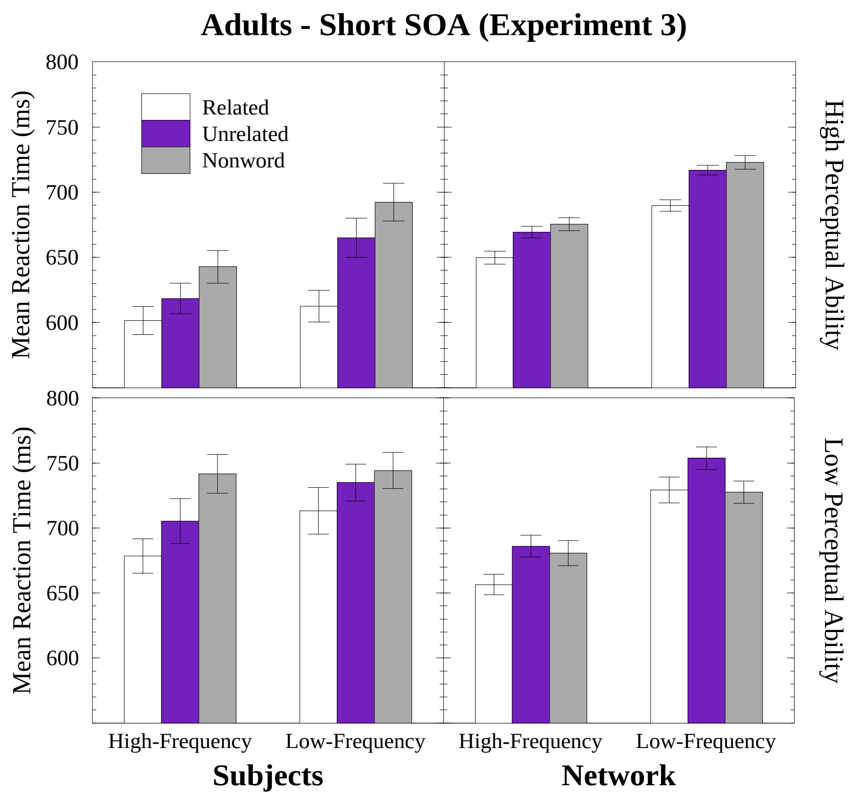

Figure 10. Mean RTs for high- and low-perceptual-ability adult subjects at the short SOA (from Experiment 3) and the mean RTs for the network from the corresponding conditions. Error bars are 1 standard error by items.

combination of frequency and perceptual ability revealed that, in the high-ability conditions, context effects were due only to facilitation (ps $<$ .001). Note that this contrasts with the corresponding conditions for adults at the long SOA, in which the simulation exhibited both facilitation and inhibition. These findings agree with the empirical results from Experiments 1 and 3 for the high-ability condition. In the low-ability conditions, context effects on high-frequency targets were also due primarily to facilitation ($p < .001$; $p = .070$ for inhibition). The effects on low-frequency targets, by contrast, were due entirely to inhibition ($p < .001$). Similar to the simulation results for Experiment 2 (children at the long SOA), the current simulation results differ from the empirical results for the low-ability condition, particularly for low-frequency targets, for which nonword primes produced faster RTs than unrelated primes.

To further clarify the pattern of facilitation and inhibition exhibited by the network, we carried out an additional analysis of the changes in these measures as a function of SOA (by collapsing the data from the simulations of Experiments 1 and 3). The magnitudes of facilitation and inhibition were normalized relative to the mean RT of the relevant priming conditions (i.e., facilitation was measured by the difference in mean RT between the nonword and related priming conditions, divided by their average; inhibition was measured by the difference between the unrelated and nonword priming conditions, divided by their average). This analysis revealed that facilitation was greater at the short compared with long SOA ($F = 9.33$, $MSE = 31.5$, $p < .005$), whereas inhibition was greater at the long compared with short SOA ($F = 29.20$, $MSE = 24.9$, $p < .001$). This general shift from facilitation at the short SOA to inhibition at the long SOA also holds numerically for the empirical data from Experiments 1 and 3 (see Figures 8 and 10) and is consistent with the findings of a number of previous studies (e.g., den Heyer et al., 1985b; Smith et al., 1987).

In summary, the network in the adult, short-SOA conditions produced a pattern of results similar to the empirical results from Experiment 3 for adults at the short SOA (200 ms), except that, for the network, the frequency-by-context interaction for the high-ability condition was only marginally reliable, and nonword primes produced overly fast RTs to low-frequency targets in the low-perceptual-ability condition.

In addition to modeling the results from Experiments 1–3, the network also provides a basis for making predictions concerning both conditions that have yet to be investigated empirically, and interactions that could not be tested with the empirical data due to the inability to control certain factors across subject groups. Specifically, the next subsection presents predictions of the performance of children at a short SOA, which was not tested empirically. The subsequent subsection presents predictions of how effects of perceptual ability, priming context, and target frequency interact with age and SOA. These interactions could not be tested reliably with the existing data because perceptual ability and target frequency were not equated across age groups, and SOA was manipulated as a between-subject factor. These predictions are important because they broaden the generality of the central empirical and theoretical claims of the current work beyond what we know from existing data.

Predictions: Children, Short SOA. A three-factor ANOVA was carried out on the network's correct mean RTs for each word in the child, short SOA conditions (70,000 word presentations; SOA = 1.0, ISI = 0.5). Figure 11 shows the means for each condition for the network; the numeric values and *SDs* can be found in Appendix 2. As was found for the other combinations of age and SOA, there were main effects of perceptual ability ($F = 60.50$, $MSE = 16{,}794$, $p < .001$), target frequency ($F = 62.25$, $MSE = 56{,}952$, $p < .001$), and priming context ($F = 103.83$, $MSE = 1819$, $p < .001$). There was also an interaction of ability and frequency ($F = 22.48$, $MSE = 16{,}794$, $p = .001$), but these effects were qualified by a three-way interaction of ability, frequency, and context ($F = 7.04$, $MSE = 1179$, $p = .001$).

When the high- and low-ability conditions were analyzed separately, both showed reliable effects of frequency and context (Fs $\geq$ 52.00, ps $<$ .001). The frequency-by-context interaction was reliable in the high-ability condition ($F = 11.95$, $MSE = 643.6$, $p < .001$), with low-frequency targets producing greater priming than high-frequency targets (ds = 44 and 22 ms, respectively). By contrast, this interaction was not reliable in the low-ability condition ($F < 1$), with equivalent priming for low- and high-frequency targets (ds = 39 and 49 ms). Note that, at least numerically, the network showed the same trend of a reverse frequency-by-context interaction in the low-ability condition as found in the other simulation conditions.

Planned comparisons of related and unrelated prime conditions with the nonword prime condition within each combination of frequency and perceptual ability indicated a pattern of facilitation and inhibition rather similar to that for the adult, short-SOA conditions (simulation of Experiment 3). Specifically, in the high-ability conditions, context effects were due only to facilitation (ps $<$ .001), whereas in the low-ability conditions, context effects were due to both facilitation and inhibition for high-frequency targets (ps $<$ .005) but only to inhibition for low-frequency targets ($p < .001$). However, given the discrepancy between the network's and subjects' performance for low-frequency targets following nonword primes in the low-ability conditions for adults as the short SOA, the prediction of inhibition for children in this con-

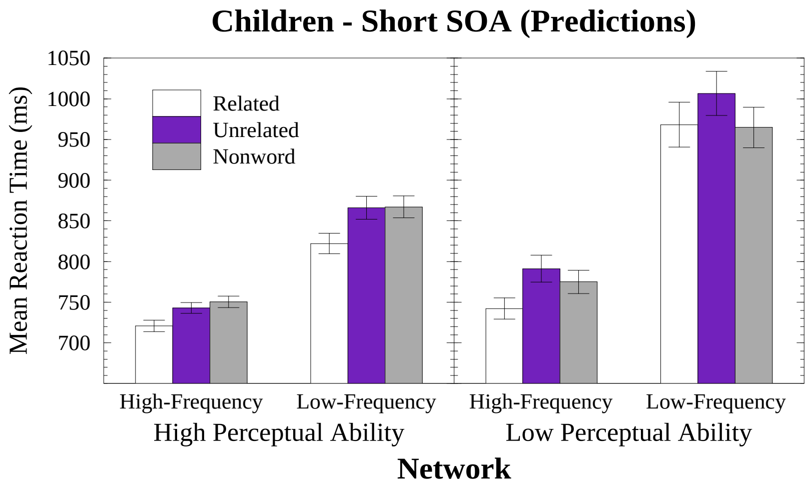

Figure 11. Mean RTs for the network under the child condition at the short SOA, as a function of perceptual ability, target frequency, and priming context. Error bars are standard errors by items.

dition must be interpreted with caution.

To determine the network's specific predictions concerning the effects of SOA on children's performance, we carried out a four-factor ANOVA on the network's correct mean RTs in the child conditions, by combining the current data at the short SOA with the corresponding data at the long SOA (from the simulation of Experiment 2). Compared with the long SOA, the short SOA produced overall faster RTs (F = 505.37, MSE = 3916, $p < .001$), smaller effects of priming context (F = 23.31, MSE = 1241, $p < .001$) and of perceptual ability (F = 75.06, MSE = 2743, $p < .001$), and a weaker context-by-ability interaction (F = 16.61, MSE = 1118, $p < .001$). Note, however, that the magnitude of the critical three-way interaction of perceptual ability, target frequency, and priming context did not depend on SOA (F = 1.82, $p > .17$).

Finally, an analysis of facilitation and inhibition effects across SOA, analogous to the one for adults reported above, was carried out for the child conditions. This analysis showed that facilitation was greater at the short compared with long SOA (F = 317.76, MSE = 48.5, $p < .001$), whereas inhibition was greater at the long compared with short SOA (F = 562.40, MSE = 40.7, $p < .001$). Thus, the network predicts that, as with adults, facilitation should diminish and inhibition should increase in moving from short to long SOA with children.

In summary, the network in the child condition at the short SOA produced a pattern of results that is generally consistent with the findings for the other combinations of age and SOA. Specifically, there was more priming for low-frequency targets compared with high-frequency targets under conditions of high perceptual ability, but equal amounts of priming for low- and high-frequency targets under conditions of low perceptual ability. Also, the high-ability conditions produced only facilitation, whereas the low-ability conditions produced a mixture of facilitation and inhibition.

Predictions: Interactions with Age and SOA. As a final analysis, correct mean RTs of the network for each word were entered into a five-factor ANOVA with age (adult vs. child), SOA (long vs. short), perceptual ability (high vs. low), and priming context (related vs. unrelated) as within-item factors, and target frequency (high vs. low) as a between-item factor. This fully crossed comparison was not possible for the empirical data because the children were not run in a short SOA condition, they were presented with a fewer number of items, and they were not matched to the adults in terms of perceptual ability. Given the increase in the power of the design relative to what would be possible in a behavioral study, and the fact that a large number of comparisons were being considered, statistical tests were considered reliable only at $p <$ .001.

Not surprisingly, the analysis of the network's RTs indicated reliable main effects of all five factors in the expected directions, and age interacted with the other four factors such that effects were larger in the child versus adult condition (ps $< .0001$). These interactions follow naturally from the asymptotic nature of the sigmoid function; the additional training for the adult condition generally results in unit activations closer to the asymptotes, thereby leaving less room for other factors to produce additional improvements in performance. Consistent with these predictions, both target frequency and priming context have been shown previously to have larger effects in poorer readers and younger children (Perfetti & Hogaboam, 1975; Schwantes, 1991; Simpson & Lorsbach, 1983; West & Stanovich, 1978).

There were two reliable three-way interactions, both involving priming context and perceptual ability. First, the effects of perceptual ability and priming context were larger at the long compared with short SOA (ps $< .0001$), but these effects were qualified by a three-way interaction between ability, context, and SOA ($F = 19.83$, $MSE = 957.9$, $p < .0001$), due to the fact that only the low-ability condition showed greater priming at the long compared with short SOA. Second, and most important, perceptual ability interacted with both priming context and target frequency (ps $< .0001$), but these effects were qualified by a three-way interaction of frequency, context, and ability ($F = 17.02$, $MSE = 2291$, $p < .005$). Specifically, the frequency-by-context interaction was reliable in the high-ability condition ($F = 30.74$, $MSE = 309.4$, $p < .0001$) but not in the low-ability condition ($F = 2.01$, $MSE = 1477$, $p > .15$). Moreover, the magnitude of this three-way interaction did not depend on age, nor on SOA, nor on their interaction (ps $> .02$). Thus, the most central finding of the current work, that greater priming for high- vs. low-frequency targets holds only for subjects with high perceptual ability, is predicted to be independent of age and SOA.

Summary of Simulation Results

The current simulation demonstrated that a distributed network model of lexical processing can account for the findings of Experiments 1–3 that priming context interacts with target frequency for subjects with high perceptual ability but not for those with low perceptual ability. As was true for the empirical studies, this finding in the network was remarkably general, holding across differences in both age and SOA.

The network also exhibited the general pattern, also true in the empirical data, that context effects resulted primarily from facilitation at the short SOA, but from a combination of facilitation and inhibition at the long SOA. This pattern is typically interpreted as implicating a combination of automatic, spreading activation at both short and long SOAs and strategic, expectancy-based processes that operate only at long SOAs. The network's behavior suggests that such a dual-mechanism account may be unnecessary, and that a single mechanism based on attractors operating over distributed representations, may provide a more parsimonious account.

General Discussion

Our empirical studies and distributed network modeling examined the influence of several theoretically important factors on the magnitude of semantic priming. Our empirical results support the extensive literature in naming and lexical decision which shows that the degree to which a prime, such as NURSE, affects the recognition of a target, such as DOCTOR, depends on the frequency of the target—low-frequency targets are influenced more by priming context than high-frequency targets (e.g., Becker, 1979; Borowsky & Besner, 1993; Stanovich & West, 1981; Stanovich et al., 1981). The most theoretically important finding of the empirical studies was that, across differences in both age and SOA, the interaction between priming context and target frequency depended on a reader's perceptual ability. Only subjects with high perceptual ability exhibited greater priming for low-frequency targets than for high-frequency targets. The subjects with low perceptual ability showed equal priming for high- and low-frequency targets (see Figures 2, 3, and 4).

We then demonstrated that a distributed network model exhibited the same pattern of results when perceptual ability was instantiated in terms of the strength with which orthographic input was presented to the network (see Figures 8, 9, and 10). In the model, the orthographic input drives the semantic representations of high-frequency targets more strongly than those of low-frequency targets, due to differences in the frequency of training on these words (also see Borowsky & Besner, 1993). As a result, the semantic system settles into a stable pattern of activity faster for high- compared with low-frequency targets. Preceding a high-frequency target by a related prime produces little facilitation relative to an unrelated prime because frequency alone is sufficient to drive the activations of semantic units near the asymptote of the sigmoid function. By contrast, priming context has a much larger effect on low-frequency targets because their initial activations fall closer to the linear range of the sigmoid, where other factors can still have clear effects (see Figure 1). Therefore, the same priming context yields a larger priming effect for low-frequency targets than for high-frequency targets due to the "diminishing returns" of the asymptotic nature of the sigmoid activation function (also see Cohen et al., 1990; Plaut, 1995; Plaut et al., 1996).

The interactive compensatory model (Stanovich, 1980) predicts that younger and low-ability readers should have equal priming effects for high- and low-frequency targets because the recognition of all words is slow and not automatic (e.g., Golinkoff & Rosinki, 1976; Guttentag & Haith, 1979; Perfetti et al., 1978; Perfetti & Hogaboam, 1975). Our distributed network model makes the same prediction under the assumption that the orthographic system of younger and low-ability readers does not strongly activate the semantic system for either high- or low-frequency targets. Because orthographic input levels for these readers fall within the linear range of the activation function for semantic units, the related and unrelated primes influence the recognition of high- and low-frequency targets to a similar degree. In other words, our model, like the interactive compensatory model, predicts that there should be no interaction between priming context and target frequency for low-ability readers, but that there should be an interaction between priming context and target frequency for high-ability readers. This is exactly what each of our empirical studies showed. Note, however, that the network model accounts for the three-way interaction between perceptual ability, priming context, and target frequency with only one mechanism, whereas alternative models invoke multiple mechanisms to account for these effects (Neely, 1977, 1991; Stanovich, 1980).

Dual-mechanism models assume that inhibition can influence word recognition only at long SOAs because it is a slow, strategic expectancy-based process, whereas facilitation influences word recognition regardless of SOA because spreading activation is fast and automatic (see Posner & Snyder, 1975; Neely, 1991). Consistent with this account, our empirical studies with adults showed both facilitation and inhibition at the long SOA, but only facilitation at short SOAs (see Figures 2 and 4).

Under conditions corresponding to adult performance (i.e., training for 200,000 word presentations), our distributed network model also exhibited inhibition dominance at a long SOA and facilitation dominance at a short SOA (see Figures 8 and 10). This finding is of fundamental importance because it is often assumed that inhibitory priming effects imply a contribution from separate expectancy-based processes. Our results indicate that the increased inhibition at long SOAs can arise from the same mechanism that produces only facilitation at short SOAs. On our account, the shift from facilitation to inhibition across SOAs reflects the degree of *hysteresis* in moving from the representation of the prime to that of the target, corresponding to the depth to which the system settles into the attractor basin for the prime when encountering the target. At a short SOA, the system has only enough time to move partially into the prime's basin. This corresponds to fairly weak semantic activity that, nonetheless, facilitates the processing of a semantically related target. By contrast, at a longer SOA, the network settles deeply into the attractor basin for the prime. On presentation of the target, the system must then move up and out of the prime's basin to derive the representation of the target. Although semantic similarity between prime and target may help this process to some extent, the semantic features for which the prime and target *differ* must nevertheless be reversed, and this process is prolonged as the prime's features (including those not shared with the target) are activated more strongly. Thus, in the adult conditions, our distributed network model provides an alternative and more parsimonious account than dual-mechanism models of the time course of facilitation and inhibition as a function of SOA. Moreover, the underlying computational principles embodied in the model are not specific to the domain of lexical processing, but apply in essentially unaltered form across the full range of cognitive processes (see McClelland, Rumelhart, & the PDP Research Group, 1986; McLeod, Plunkett, & Rolls, in press; Quinlan, 1991).

Dual-mechanism models also predict that children should exhibit inhibition as well as facilitation because their word recognition processes are slow and not automatic (Stanovich, 1980). Our empirical results did not support this prediction—both high- and low-perceptual-ability children exhibited facilitation but no inhibition (see Figure 3). It must be acknowledged, though, that our distributed network model also provided a less than adequate account of these findings (see Figure 9). Specifically, the model showed clear inhibitory effects for the low-perceptual-ability conditions that were absent in the empirical data for the subjects. We will address this discrepancy below, in the context of considering the full range of limitations of the model. Following this, we discuss other empirical findings which may appear to challenge our account, and describe empirical and computational extensions of our approach to address related phenomena. We start, though, by articulating the specific empirical predictions made by our model.

Empirical Predictions of the Model

An important benefit of developing an explicit computational implementation that instantiates the core principles of a theory is that it can be used to generate specific quantitative predictions of performance under novel conditions. A case in point concerns the predictions of the network for the performance of children when tested at a short SOA. The network's RTs were scaled to milliseconds by regressing the network's mean settling times against the mean RTs of subjects across the 36 conditions in the current empirical studies. Applying the same regression equation to the network's settling times when tested in the child condition at the short SOA yields specific predictions of mean RTs in milliseconds across 12 additional conditions (3 priming contexts $\times$ 2 target frequencies $\times$ 2 levels of perceptual ability; see Figure 11). Moreover, by comparing these results to those obtained in the child condition at the long SOA, specific effects of SOA on performance are predicted. In particular, children tested at the short SOA should produce faster RTs and smaller effects of priming context and perceptual ability than at the long SOA, as we found for adults. Most important, the model predicts that children at a short SOA should exhibit the same interaction between perceptual ability, priming context, and target frequency shown by adults at both the long and short SOA and by children at the long SOA.

The second prediction generated by the network model was that the interaction between perceptual ability, target frequency, and priming context (when defined as related versus unrelated) was independent of age and SOA. Unfortunately, the empirical data from the current set of studies cannot be used to evaluate this prediction because the adults and children were not given the same list of prime-target pairs, and because the perceptual ability of the adults differed between the long- and short-SOA experiments. Future research could address these issues by manipulating SOA as a within-subject factor, and by administering the same prime-target pairs to children and adults. When facilitation and inhibition effects were examined relative to a nonword prime baseline, the model predicted a general shift from facilitation at the short SOA to inhibition at the long SOA, for both adults and children. With respect to adults, these predictions are broadly consistent with the past literature, which shows no increase in facilitation but an increase in inhibition with longer SOAs (den Heyer et al., 1985b; Smith et al., 1987). Again, this has not been tested in children.

Finally, although more tentatively, the network produced a trend toward a reverse frequency-by-context interaction (i.e., greater priming for high- compared with low-frequency targets) when tested in the adult, low-perceptual-ability condition at the long SOA. The same pattern held numerically for these conditions at the short SOA. The corresponding empirical data showed the same direction of effects numerically although these were not statistically reliable. As discussed in the context of the simulation of Experiment 1, this reverse interaction, like the standard one in the high-ability condition, can be understood in terms of the nonlinear effects of the sigmoid activation function (see Figure 1). Specifically, if the input to semantic units under the low-ability condition is sufficiently weak to begin to encroach onto the opposite side of the sigmoid for low-frequency targets, the effects of priming context will be reduced relative to that for high-frequency targets, which remain within the linear portion of the sigmoid (but see Footnote 18). Given that a reverse frequency-by-context interaction held only weakly in the model, however, additional simulation work is needed to verify that this effect is indeed a robust prediction of the model.

Limitations of the Model

Empirical Adequacy. The primary limitation of the implemented model in accounting for the findings of the empirical studies concerns its exaggerated inhibition in low-perceptual-ability conditions, particularly in the child conditions (as mentioned earlier) but also for low-frequency targets in the adult conditions. There are two factors which may have contributed to this discrepancy.[19] The first concerns the adequacy of the nonwords used as a neutral baseline priming context. Specifically, the nonwords had a VCV orthographic structure, whereas all of the orthographic patterns experienced during training (i.e., words) had a CVC structure. Although there are clear differences in the orthographic structure of the words and nonwords used in the empirical studies (see Appendix 1), these differences are not as extreme as in the simulation. One consequence of this may have been that the nonwords did not engage the network sufficiently strongly, particularly in the low-perceptual-ability conditions where the input strength was relatively weak. As a result, compared with unrelated word primes, nonword primes produced far less hysteresis, and hence much faster RTs, for subsequent target words, leading to an exaggeration of inhibitory effects in the network. In essence, our VCV nonwords are subject, to some extent, to the same criticism leveled against neutral baselines like Xs and words like READY—such repeated stimuli may not have the same attentional effects or engage the same levels of linguistic processing as word or word-like nonword primes (see Antos, 1979; Jonides & Mack, 1984; McNamara, 1994; McKoon & Ratcliff, 1992; Neely, 1991, for discussion).

The second factor that may have increased the inhibition shown by the model compared with the subjects concerns the instantiation of semantic and associative relatedness in the model. The prime-target stimuli for our empirical studies were chosen from free-association norms (Nelson et al., 1994). Although some of these pairs were in the same semantic category, many were not. By contrast, in the model, all associatively related prime-target pairs were also categorically related—there were no purely associatively related pairs. Empirical studies show that inhibition effects are larger in categorical priming than in associative priming (Lupker, 1984), and Plaut (1995) observed the same tendency in a simulation very similar to the current model but which separated categorical and associative relatedness. Thus, the greater predominance of categorical relatedness among prime-target pairs for the model compared with the subjects may have contributed to the overly strong inhibitory effects in the former.

Finally, the model produced a stronger frequency effect than was observed in the empirical data, particularly in the child conditions (see Figure 9). Moreover, for both the adult and child conditions at the long SOA, the frequency effect interacted with perceptual ability, being larger in the high- versus low-ability conditions. This finding agrees with previous empirical data on ability differences in children (Perfetti & Hogaboam, 1975); however, our empirical studies exhibited only trends in the same direction. As previously suggested, this discrepancy may due to the fact that the frequency manipulation in the network model was

[19] In addition, it is possible that the greater proportion of word versus nonword targets following word primes may have led to an underestimate of inhibition relative to facilitation in the empirical studies. However, we have argued that this effect is likely to be negligible in the relevant conditions—namely, children at the long SOA (Experiment 2) and adults at the short SOA (Experiment 3).

dichotomous—high-frequency words were presented exactly twice as often as low-frequency words—whereas the frequency manipulation in the empirical studies was continuous (based on Kučera & Francis, 1967). If the empirical manipulation of frequency were stronger, we would expect a significant interaction between perceptual ability and target frequency. Also note that the semantic priming effects in the network, like the frequency effect, is larger than those in the human subjects. Again, this may have been because the relatedness manipulation in the network was strong and dichotomous, whereas the manipulation in the empirical studies was a continuous variation in associative strength (Nelson et al., 1994).

Overall, while there are certainly aspects of the performance of our distributed network model which are inconsistent with the findings from our empirical studies, these differences can be understood in terms of the simplifications and approximations made in developing the model. We leave it to future research to determine whether a more realistic model can provide an even closer fit to the empirical data.

Simplifications. In addition to the simplifications just mentioned, three others need to be considered more carefully. The first concerns the means by which reaction times and lexical decisions are determined. In the current implementation, the RT of the network to a stimulus is determined by the number of processing cycles required for semantic activation levels to stabilize below some specified criterion. At that point, a lexical decision is made on the basis of the degree to which semantic representations have been driven strongly towards binary activation levels, operationalized by a measure termed stress (see Equation 4 and Figure 6), such that a stress level above a particular criterion indicates a "yes" response (also see Plaut, 1997).

Stress provides a reliable basis for discriminating words from nonwords due to the lack of systematicity between orthography and semantics. Words generate high stress levels because they are trained to generate semantic representations consisting of binary features. Nonwords, by definition, were not presented during training. The network's behavior for these items is solely a function of generalization from its knowledge of orthographically similar words. Given that all of the network's knowledge is encoded in the same set of connection weights, processing a nonword partially engages the mappings for all of the trained words, in proportion to the orthographic similarity of each word to the nonword. For a systematic mapping, like that between orthography and phonology, the mappings for orthographically similar words generally agree with each other and thus conspire effectively to generate strong output activation for nonwords (see Plaut et al., 1996). By contrast, for an unsystematic mapping, like that between orthography and semantics, orthographically similar words map to unrelated sets of semantic features. Nonwords still engage a combination of the mappings for similar words, but now these mappings conflict with each other—semantic units that are activated by the mapping for one similar word are inhibited by the mappings for different similar words. As a result of this inconsistency, semantic activations are driven less strongly towards extreme values—yielding lower stress—when the network processes a nonword compared with when it processes a familiar word. Although the current model used only a relatively small number of words (128), Plaut (1997) showed that semantic stress can support accurate lexical decision for 2998 words (Plaut et al., 1996).

A number of researchers (e.g., Besner & Joordens, 1995; Borowsky & Masson, 1996; Joordens & Becker, 1997; Masson & Borowsky, 1995; Rueckl, 1995) have recently proposed a related measure—the negative of *energy* (Hopfield, 1982), sometimes termed *harmony* (Smolensky, 1986) or *goodness* (McClelland & Rumelhart, 1988)—as the basis on which subjects make lexical decisions. One drawback of this measure, $-\sum_{i<j} a_i a_j w_{ij}$, is that it requires decision processes to have direct access to the weights, w_{ij}, among units in the lexical system. By contrast, computing the stress measure requires only a fairly simple combination of unit activations. It should be pointed out, though, that the two measures are closely related—as long as all output activations are on the correct side of "neutral" (i.e., 0.5 for standard [0,1] units), then increasing stress by moving activations towards more binary values will generally also decrease energy.

Regardless of whether stress or energy is used to make lexical decisions, however, there is still a problem with the procedure employed in the current simulation, stemming from the use of separate criteria for determining *when* the network responds (stability) and *how* it responds (stress). A more satisfactory approach would be to define a response criterion based on how stress values change over the course of settling in response to word and nonword inputs (see Joordens & Becker, 1997, for a related proposal in terms of "harmony"). We employed the simpler procedure of defining RTs in terms of settling times partly because this approach has been used previously to model response latency data from lexical tasks successfully (Becker et al., 1997; Borowsky & Masson, 1996; Kawamoto, 1993; Kawamoto, Farrar, & Kello, 1994; McRae et al., 1997; Masson, 1995; Plaut et al., 1996), but also because implementing the actual mechanism that generates "yes" and "no" responses was considered beyond the scope of the current work.

It should be noted, though, that decision processes can be modeled effectively using the same computational principles as employed in the current work. For example, Usher and McClelland (1995) have demonstrated recently that competition between linear, stochastic, time-averaging units representing alternative responses gives rise to a number of basic properties of empirical findings in standard choice RT tasks. Their approach could be applied in the current context by adding to the lexical network a response layer consisting of two units (see Figure 12a): a "yes" unit whose input is the level of semantic stress for the current target, and a "no" unit whose input is the value of the decision criterion used in the current work to distinguish words from nonwords (which could be estimated from a running average of stress values across all stimuli). The "yes" and "no" units compete based on the relative strength of their inputs, and the network responds when the activation of one of the units exceeds a threshold response criterion. An important property of this type of competitive response system is that the competition takes longer to resolve, and hence RTs are prolonged, when the inputs to the response units are of similar magnitude (i.e., when a word or nonword target produces a stress value close to the decision criterion separating the word and nonword distributions). Moreover, under experimental conditions in which the separation of the word and nonword distributions is increased, the competition generally resolves more quickly, and thus a more aggressive response criterion (i.e., lower activation threshold) can be used to speed overall responding while keeping error rates acceptably low (see Figure 12b). These characteristics play an important role in our accounts of stimulus blocking effects, as discussed below.

The second and perhaps more obvious simplification in the current implementation was that it did not include phonological representations and processes. This omission might seem particularly problematic in light of recent findings of very rapid phonological influences on lexical processing (Booth, Perfetti, & MacWhinney, submitted; Lukatela, 1994; Lukatela, Savic, Urosević, & Turvey, 1997; Lukatela & Turvey, 1990, 1991, 1994; Perfetti & Bell, 1991; Van Orden, 1987; Van Orden, Johnston, & Hale, 1988, although see Jared & Seidenberg, 1991; Verstaen, Humphreys, Olson, & d'Ydewalle, 1995). In fact, other network models and empirical findings have illustrated the importance of examining the interaction among orthographic, phonological, and semantic representations when trying to account for behavioral data from naming and lexical decision tasks (see, e.g., Kawamoto, 1993; Plaut, 1997; Plaut et al., 1996; Stone, Vanhoy, & Van Orden, 1997; Strain et al., 1995; Van Orden & Goldinger, 1994; Van Orden, Pennington, & Stone, 1990). Thus, the current simulation, which involved only a mapping from orthography to semantics, cannot be expected to provide a full account of lexical processing in general, nor even of lexical decision performance in particular.

Nonetheless, apart from considerations of issues that relate to phonology *per se*, such as pseudohomophone effects (McCann, Besner, & Davelaar, 1988), the central properties exhibited by the current implementation would also be expected to hold for a more general implementation that included phonology. The reason is that, in English, orthography and phonology bear a similar relationship to semantics—the similarity of monomorphemic words within each domain is essentially unrelated to their semantic similarity. As explained above, it is this lack of structure between the surface forms of words and their meanings that, on the current account, provides the most reliable basis for distinguishing words from nonwords. Thus, the rapid derivation of phonological information from orthography allows two unstructured mappings to contribute to performance instead of one, but does not fundamentally alter the relative effectiveness of familiar versus novel surface forms (i.e., words vs. nonwords) to engage semantics. Of course, pseudohomophones (e.g, BRANE) are precisely those stimuli which violate the more general functional similarity of orthography and phonology, because they are orthographically unfamiliar but phonologically familiar. For exactly this reason, pseudohomophone effects in lexical decision are beyond the scope of the current implementation (but not, of course, outside the scope of the more general theoretical framework of distributed network models; see, e.g., Plaut, 1997; Seidenberg & McClelland, 1989).

Finally, the current model bases lexical decisions on stress calculated only over semantic representations. While there is strong evidence that semantics plays an important role in lexical decision performance (see, e.g., Azuma & Van Orden, 1997; Balota & Chumbley, 1984; Balota, Ferraro, & Connor, 1991; Besner & Smith, 1992; Borowsky & Besner, 1993; Borowsky & Masson, 1996; Chumbley & Balota, 1984; Hino & Lupker, 1996; James, 1975; Jastrzembski, 1981; Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989), it is also clear that readers can base lexical decisions, at least in part, on orthographic and/or phonological information, particularly when the nonword foils are relatively unwordlike (James, 1975; Waters & Seidenberg, 1985). It is important to note that our nonword foils were, for the most part, orthographically legal, so reliance solely on orthography is unlikely. In a more comprehensive version of our account of lexical decision, we would assume that subjects can base their decisions on any available information in the lexical system, and that they adopt a strategy that optimizes their performance given the composition of the stimuli (also see Seidenberg & McClelland, 1989). Moreover, given that orthographic information is available earlier than either phonological or semantic information, we would expect subjects to rely on orthographic information to whatever extent possible. Given our focus on semantics rather than orthography as a basis for lexical decision, we would not expect the current form of our model to account for the

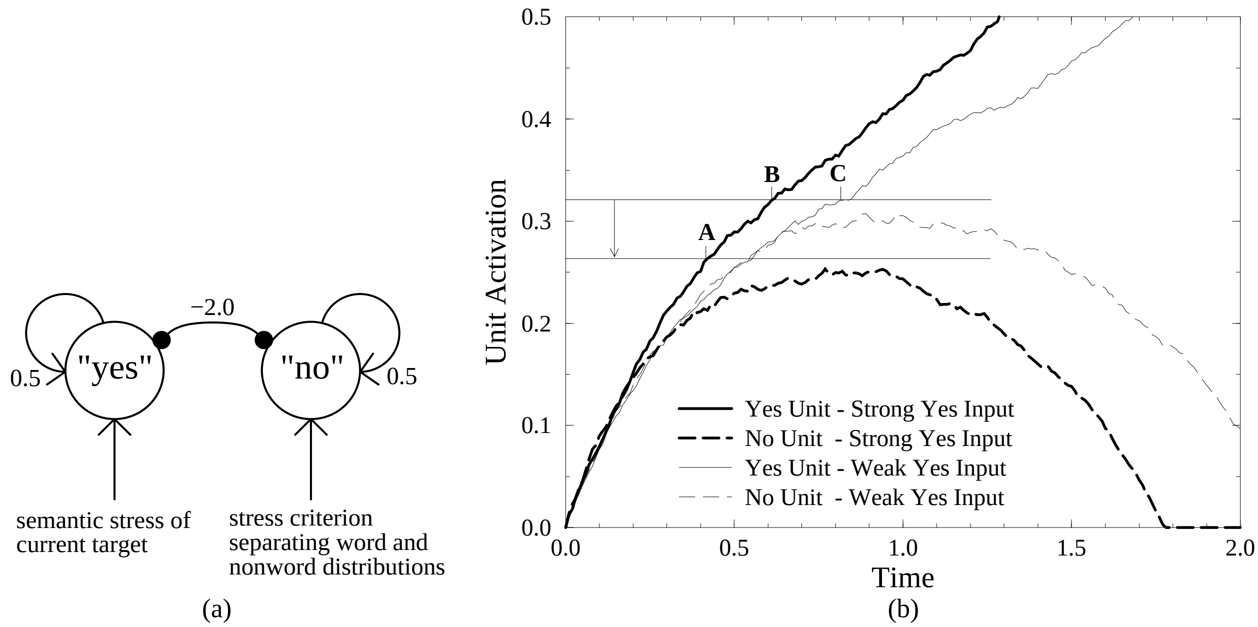

Figure 12. (a) A depiction of competitive "yes" and "no" response units (Usher & McClelland, 1995) and (b) their activations when the input to the "yes" unit is relatively strong and much larger than the input to the "no" unit (i.e., 0.9 vs. 0.8, respectively) compared with when it is weaker and more similar to the "no" input (i.e., 0.85 vs. 0.8, respectively). These input values are intended to reflect the semantic stress for a target word ("yes" inputs) and for the stress criterion separating word and nonword responses ("no" inputs). In addition to these inputs, each unit has an excitatory weight of 0.5 from itself, an inhibitory weight of -2.0 from the other unit, and its activation is corrupted with noise (SD = 0.2) and integrated with time constant τ = 0.01. The horizontal lines indicate alternative response criteria. Note that, when strong and weak word inputs are mixed, responses to the former are faster (cf. B vs. C). When the strong word inputs are blocked, the response criterion can be shifted downward (indicated by the arrow) to yield even faster RTs for these items (cf. A vs. B).

influence of orthographic factors, such as neighborhood density, on performance (see, e.g., Andrews, 1992; Sears, Hino, & Lupker, 1995).

In summary, there are a number of ways in which the current implemented model falls short of a fully comprehensive account of lexical processing. Nonetheless, despite the simplifications incorporated into the model, we claim that the central computational principles underlying its performance can be extended to account for the full range of relevant phenomena.

Additional Empirical Issues

Having discussed limitations of our implemented model of semantic priming in lexical decision, we now turn to a consideration of related empirical phenomena which would seem to challenge our more general single-mechanism, distributed network account. These phenomena relate to blocking and strategy effects, priming across unrelated items, categorical versus associative priming, and backward associative priming. We devote considerable attention to blocking and strategy effects because these would seem to be the most problematic for our account.

Blocking and Strategy Effects. Perhaps the most difficult challenge to a single-mechanism account of lexical processing concerns the effects of experimental manipulations that induce apparent changes in the processing strategies adopted by subjects. It is often claimed that single-mechanism models cannot account for such strategic effects in visual word recognition (see, e.g., Neely, 1991). These effects have been operationalized in a number of ways. The most common way is by manipulating properties of the experimental stimuli—for example, the similarity of nonword foils to words (Stone & Van Orden, 1993), the proportion of prime-target pairs that are related (Neely, 1977), or the proportion of pairs that are associatively versus categorically related (Becker, 1980). For example, increasing the proportion of related trials yields larger semantic priming effects at long but not short SOAs (de Groot, 1984; den Heyer, 1985; den Heyer, Briand, & Dannenbring, 1985a; Neely, 1977; Neely, Keefe, & Ross, 1989; Seidenberg, Waters, Sanders, & Langer, 1984b; Tweedy, Lapinski, & Schvaneveldt, 1977). Another type of manipulation is to compare the effects of blocking versus mixing experimental conditions for the same stimuli. For example, Smith, Besner, and Miyoshi (1994) found that the magnitude of semantic priming at a short SOA depended on whether such trials were blocked or mixed with long SOA trials; adult subjects exhibited priming at the short SOA in the blocked condition, but such priming effects were minimal in the mixed condition. This suggests that semantic priming is at least partially subject to strategic effects and, therefore, is not entirely automatic. Consistent with this conclusion, we have preliminary evidence that, if short and long SOAs are mixed, the interaction between perceptual ability, priming context, and target frequency holds only at a long SOA (Plaut & Booth, in preparation).

It is clear, then, that the nature of processing a given target word depends not only on the immediately preceding (priming) context but also on more general aspects of the experimental situation. The critical question is what sort of processes must be postulated to account for these types of strategic effects, and to what extent are these specific to the lexical system per se. Blocking and list context effects are typically explained by reference to changes in the operation of expectancy-based processes (see Becker, 1980, 1985; Neely, 1991; Neely & Keefe, 1989). The fact that such processes are assumed to be slow accounts for the fact that many strategic effects arise only at relatively long SOAs (but see Smith et al., 1994; Stolz & Besner, 1997; Stolz & Neely, 1995).

An alternative approach to explaining these types of effects, however, is to assume that subjects adjust the operation of decision processes, which are separate from the lexical system *per se*, as a function of the composition of the stimuli and testing conditions within the current block (see, e.g., Gordon, 1983; Grainger & Jacobs, 1996; Seidenberg et al., 1984b; Stone & Van Orden, 1993; Waters & Seidenberg, 1985). In particular, as suggested earlier, subjects might adjust the response criterion within a competitive response system (Usher & McClelland, 1995, see Figure 12a).[20] Recently, Plaut (1997, also see Gordon, 1983; Seidenberg & McClelland, 1989) has argued that, if lexical decisions are based on semantic stress, this approach can provide an account of the frequency blocking effect (Glanzer & Ehrenreich, 1979), in which RTs to high- but not low-frequency targets are reduced under blocked compared with mixed presentation. The essence of the account is that, because high-frequency words tend to produce higher stress values than low-frequency words, a more conservative response criterion is needed to produce acceptable levels of accuracy when low-frequency words are among the stimuli (whether blocked or mixed). However, when high-frequency words are blocked, there is greater separation between the word and nonword distributions, so that a more aggressive response criterion can be adopted that, for the same error rate, produces faster responding (see Figure 12b). The same form of account can explain why the effects of target frequency are increased

[20]Similarly, with regard to the naming task, Lupker, Brown, and Colombo (1997) and Jared (1997) have recently provided evidence that shifts in a criterion for when to initiate articulation provides a better account of blocking effects in naming (e.g., Monsell, Patterson, Graham, Hughes, & Milroy, 1992; Paap & Noel, 1991) than accounts which rely on changes in the relative contribution of lexical and sublexical pathways. Note, though, that Lupker and colleagues propose a time-based criterion, whereas the current proposal involves an activation-based criterion.

when nonword foils are more wordlike (Stone & Van Orden, 1993).

An analogous account may explain the SOA blocking effects (Smith et al., 1994; Stolz & Besner, 1997).[21] In our model, RTs are faster and priming effects are weaker at the short SOA because, compared with the long SOA, the network does not settle as deeply into the prime's attractor basin, so there is less hysteresis in moving from the pattern produced by the prime to the representation of the target. Moreover, given that the network is settling to a binary semantic pattern, the fact that the network settles faster at the short SOA implies that, at any point in time following the presentation of the target, the value of semantic stress is generally higher at the short versus long SOA. By contrast, the stress levels for nonword targets are relatively unaffected by manipulation of SOA. Consequently, there is greater separation between the distributions of stress values for words and nonwords when short-SOA trials are blocked compared with when they are mixed with long-SOA trials. Following the logic used in explaining the frequency blocking effect, a more conservative response criterion is required in the mixed versus blocked condition for short SOA trials. Under this more conservative criterion, the relative difference in stress for targets following related versus unrelated primes is reduced (compared to the blocked condition), thereby reducing and perhaps even eliminating priming effects at the short SOA in the mixed condition. By contrast, essentially the same response criterion is needed for long-SOA trials whether they are mixed or blocked, so there is no effect of this manipulation on the magnitude of semantic priming at long SOAs.

Now consider the findings that increasing the proportion of related trials produces stronger semantic priming (de Groot, 1984; den Heyer, 1985; den Heyer et al., 1985a; Neely, 1977; Neely et al., 1989; Seidenberg et al., 1984b; Tweedy et al., 1977). Note that the basic semantic priming effect is that RTs to targets are faster following related compared with unrelated primes. Based on the argument just presented for short versus long SOAs, this means that, at a given point in processing the target, stress values are generally higher following related versus unrelated primes. Increasing the proportion of related trials is, thus, analogous to blocking short-SOA trials: the manipulation permits a more aggressive response criterion which increases the relative difference in stress values for the related versus unrelated priming conditions, thereby leading to a larger priming effect. Relatedness proportion may not affect semantic priming at short SOAs (de Groot, 1984; den Heyer et al., 1985a; Neely, 1977) because, compared with long SOAs, the basic priming effects are weaker at short SOAs (see Neely, 1991) and, thus, less susceptible to modulation.

Finally, adjustment of a response criterion may also account for Becker's (1980) finding that categorically related prime-target pairs produce inhibition dominance in the context of other categorically related pairs, but facilitation dominance in the context of predominantly associatively related pairs. An examination of Becker's data suggests two contributing factors. The first relates to the fact that, in the blocked condition (Experiment 4), RTs are much faster for associatively related pairs (antonyms) than for categorically related pairs (category name to exemplar). In the mixed condition (Experiment 5), RTs to both types of pairs are slower, but much more so for the associatively compared with categorically related pairs. This corresponds to the pattern predicted by a shift to a more conservative response criterion in the mixed versus blocked conditions, causing greater slowing for the faster condition—the associatively related pairs. The second contributing factor is that, for trials involving related or unrelated primes, the part-of-speech of the target was almost always the same as that of the prime—antonyms are typically adjectives whereas category names and exemplars are typically nouns.[22] This predictability may have facilitated subjects' responses in both of these conditions relative to the neutral priming condition (i.e., a string of Xs). The added facilitation for the related and unrelated conditions would be expected to have its greatest impact when performance is otherwise the slowest—in the mixed condition for categorically related pairs—which is exactly the condition which yielded the paradoxical facilitation dominance. By contrast, the facilitation is not as strong in the blocked condition, and so the categorically related pairs continue to show the standard pattern of inhibition dominance.

To be clear, the proposals we have outlined here are not fully adequate accounts of the relevant phenomena; rather, they are intended to sketch out an approach to explaining these effects which relies only on the adjustment of a response criterion, rather than on the existence of complicated expectancy-based processes. It seems unlikely, however, that changes in a response criterion will provide an account of other types of strategic effects—

[21]Smith et al. (1994) provide an explanation for the SOA blocking effect which they and Stolz and Besner (1997) describe as a "signal-detection" account. This account is, however, rather different than the current proposal, in that it postulates the adjustment of a criterion that controls whether lexical activation spreads from the orthographic system to the semantic system. By contrast, the relevant criterion on the current account is an activation threshold applied to "yes" and "no" units within a competitive response system (Usher & McClelland, 1995) which is not considered part of the lexical system *per se*.

[22]To avoid a part-of-speech confound in comparing associatively related versus categorically related priming, Becker (1980) included fillers that were associatively related pairs in which the targets could be either nouns or adjectives. This manipulation does not, however, prevent primes and targets from having the same part-of-speech, which is generally true of strong associates (e.g., see Appendix 1). Unfortunately, Becker did not provide his stimuli in reporting his findings.

particularly those involving changes in the instructions to subjects (Favreau & Segalowitz, 1983; Neely, 1977) and other depth-of-processing manipulations (e.g., Henik, Freidrich, & Kellogg, 1983; Kaye & Brown, 1985; Smith, Theodor, & Franklin, 1983). Of course, instructions must induce strategic effects on lexical processing at some level; otherwise, how can it be that, when presented with the same stimulus, subjects perform lexical decision in one experiment and naming or letter search in others? Thus, the relevant question is not whether higher-level strategies influence processing, but rather whether the lexical system itself must incorporate a strategic mechanism such as Becker's (1980) generation of expectancy sets or Neely and Keefe's (1989) retrospective semantic matching. Our position is that the mechanisms that underlie these types of strategy changes apply generally across all cognitive domains and are not specific to the lexical system, although they can certainly influence processing in this domain. In fact, on the current account, even the response criterion is not part of the lexical system, but is part of a more general cognitive mechanism for making forced-choice decisions (Ratcliff, 1978; Usher & McClelland, 1995). Thus, although we do not, at present, have a fully adequate theory of the nature and operation of these general-purpose mechanisms, we can nonetheless make progress in articulating the principles of operation of the lexical system quite apart from such a theory. In this way, our model of the lexical system per se remains a single-mechanism account.

Priming Across Unrelated Items. Another apparent challenge for distributed network models is to account for the finding that associative priming can span an intervening item, such as in the word sequence NURSE–CAT–DOCTOR (e.g., Joordens & Besner, 1992; McNamara, 1992; Meyer & Schvaneveldt, 1971). Although these priming effects are weak, they have challenged distributed network models because if the network settles completely to the meaning of the intervening word CAT, then the pattern of activity representing the meaning of NURSE will be completely eliminated, leaving no opportunity for it to facilitate the processing of DOCTOR. However, the intervening word might be processed only partially, leaving residual semantic activation for NURSE to influence processing of DOCTOR (Masson, 1995). Indeed, Plaut (1995) showed that a distributed network model exhibited associative priming across an intervening item, particularly under conditions which encourage fast responding (also see Masson, 1995). Therefore, distributed network models seem able to account for the existence of priming across an intervening item without recourse to another mechanism.

Categorical Versus Associative Priming. A model of semantic priming should also account for the time course of facilitation and inhibition in categorical versus associative priming. At short SOAs, there is facilitation dominance for both categorical and associative priming, whereas at long SOAs there is facilitation dominance for associative priming, but inhibition dominance for categorical priming (den Heyer et al., 1985b; Smith et al., 1987). Plaut (1995) showed that a distributed network model exhibited greater associative priming with longer SOAs, and that same model also exhibited a decrease in categorical priming from short to long SOAs. Although this model replicated the basic finding in the literature, conclusions could not be drawn about the relative magnitudes of facilitation and inhibition because the performance of the model was not evaluated relative to a neutral baseline condition. Although the current simulation did employ a neutral (nonword) priming baseline, it involved a training environment with complete co-occurrence of categorical and associative relatedness and, thus, it cannot be used to evaluate the relative contributions of these factors.

However, a version of our simulation that at least partially separated categorical and associative relatedness might be able to account for the time course of these types of priming. Associative priming occurs in our model because, during training, the network learned to make a rapid transition from the representation of a prime to that of a target. This learned transition produces strong facilitation which increases with SOA because the representation generated by the prime becomes increasingly accurate. Note, however, that actual words typically have only one or at most a few strong associates; thus, the majority of words that precede a given target word during training are unrelated to it. As a result, there is minimal inhibition from unrelated primes because the network has learned to ignore nonoverlapping previous patterns except in the few specific cases involving associative relatedness.

On the other hand, categorical priming occurs in a distributed network model because a related prime activates features that overlap with those of the target. Categorical facilitation tends to be weak because only some features overlap between the prime and target, whereas categorical inhibition tends to be strong because many features do not overlap. In fact, this may explain why some authors have failed to find priming for categorically but not associatively related prime-target pairs (Shelton & Martin, 1992; Moss et al., 1995) and why others have found pure categorical priming only when the prime-target pairs are very highly related (Lund, Burgess, & Atchley, 1995; McRae & Boisvert, in press; Perea & Gotor, 1997). Inhibition increases with longer SOAs because, with additional processing, semantic units that differ between the prime and target are driven to more extreme values. In order to correctly identify the target, all of these differences must be corrected, so the magnitude of inhibition is greater at longer SOAs.

Backward Associative Priming. A model of visual word recognition should also be able to account for the existence of backward priming in lexical decision, and the absence of this effect in naming (Seidenberg et al., 1984b). In backward priming, the prime and target are related only through a backward association from target to prime (PAN–BED) and not through a forward association from prime to target. One of the greatest triumphs of compound-cue theory (Ratcliff & McKoon, 1988) was that it could account for backward priming by assuming that subjects use a familiarity value of the prime-target combination in order to make lexical decisions to the target. Compound-cue theory has not been extended to account for the absence of backward priming in naming. Can distributed network models account for these paradigm differences in backward priming?

The current simulation cannot be used to test for the existence of backward priming because, as just pointed out, associative relatedness always co-occurred with categorical relatedness and the latter involves a symmetric relationship. However, Plaut (1995) failed to find backward priming in a network in which associated prime-target pairs were not also categorically related. Thus, it seems unlikely that backward associative priming is an inherent property of distributed network models.

There is, however, a version of the compound-cue account of backward associative priming that is consistent with distributed models. This account relies on the fact that, as far as we know, backward priming has been demonstrated only with English stimuli in which a high proportion of associates form compound words (e.g., BED–PAN; Peterson & Simpson, 1989; Seidenberg et al., 1984b; Shelton & Martin, 1992). In English, the type of a compound's referent is determined by the second component (i.e., a BEDPAN is a type of PAN; Marchand, 1969). Thus, the representation generated by the prime PAN would be expected to overlap with that of BEDPAN due to a categorical relationship. This representation, in turn, would facilitate the processing of BED due either to additional featural overlap (e.g., a functional relationship—a BEDPAN is a pan used in BED; Moss et al., 1995) or perhaps to top-down activation of orthography from semantics (also see Borowsky & Besner, 1993). Note that the backward priming effect is generally much weaker than the standard semantic priming effect (see Neely, 1991). The fact that backward priming has been observed in lexical decision but not in naming may be due to the fact that, on the current account, lexical decision is based at least in part on semantic activation whereas naming is based on phonological activation. As elaborated below, due to the high degree of systematicity between English orthography and phonology, orthography may drive phonology so strongly that semantics has little opportunity to influence naming performance. As a result, standard semantic priming is weaker in naming than in lexical decision (Keefe & Neely, 1990; Lorch et al., 1986; Lupker, 1984; Neely et al., 1989); backward priming, which is already weak in lexical decision, may become too small to measure in naming.

Extensions of the Approach

The most obvious extension of the current work, both empirically and computationally, would be to address semantic priming effects in naming. As just mentioned, semantic priming is weaker in naming than in lexical decision, often being present only in a subset of the conditions that produce semantic priming in lexical decision (see Neely, 1991). One possible exception to this pattern is mediated priming (i.e., LION–STRIPES, presumably via TIGER), which has been observed in naming but not in lexical decision when tested under comparable conditions (Balota & Lorch, 1986, but see McKoon & Ratcliff, 1992; McNamara, 1992; McNamara & Altarriba, 1988; Shelton & Martin, 1992).

Semantic priming effects in naming could be addressed within a distributed network model in which orthographic, phonological, and semantic representations interact to settle simultaneously on the appropriate meaning and pronunciation for written words (Kawamoto, 1993; Plaut et al., 1996; Seidenberg & McClelland, 1989). The current approach to modeling semantic priming in lexical decision could be extended directly to the naming task by basing responses on phonological rather than semantic activation, and by allowing residual activation from a previous stimulus to influence the processing of a current target word.

An important aspect of the English lexical system is that there is a high degree of systematicity between orthography and phonology, but essentially no systematicity between either of these and semantics (at a monomorphemic level). A fundamental property of distributed networks models is that they learn systematic mappings more quickly and strongly than unsystematic mappings (Kawamoto, 1993; Plaut et al., 1996; Van Orden & Goldinger, 1994; Van Orden et al., 1990). As a result, while orthographic input will activate both phonological and semantic representations simultaneously, the phonological representations will settle far more quickly and will be less sensitive to pre-existing activation (see, e.g., Kawamoto, 1993; Kawamoto & Zemblige, 1992). These properties provide a natural account of why semantic priming effects are weaker in naming (based on phonological activation) than in lexical decision (based on semantic activation). Consistent with this account, naming performance can exhibit similar semantic priming effects to that found in lexical decision if it is slowed to a comparable rate by experimental manipulations such as target masking (Flores d'Arcais, Schreuder, & Glazenborg, 1985) or lateralized presentation (Chiarello, Burgess, Richards, & Pollock, 1990).

Two predictions concerning semantic priming in naming arise directly from the effects in a distributed network model of a greater degree of systematicity between orthography and phonology than between orthography and semantics. First, following the previous discussion, the effects of variables such perceptual ability and target frequency should be weaker in naming than in lexical decision. Second, given that generating the semantic representation of the prime is relatively slow compared with generating its phonology, semantic priming effects in naming should be larger at long SOAs than at short SOAs. We have conducted a naming experiment to test these two predictions in adult subjects (Plaut & Booth, in preparation). In support of these predictions, the effect of priming context was reliable only at the long SOA, and was weaker overall than in the current lexical decision experiments. Furthermore, the priming effects at the long SOA were equal for both high- and low-perceptual-ability subjects and for both high- and low-frequency words. These results contrast with the findings of the current lexical decision experiments, in which semantic priming effects depended on target frequency as well as on perceptual ability.

Our failure to find differential priming effects for high- versus low-frequency targets in naming may, in part, be due to the fact that, in English, frequency effects are in large part restricted to words with inconsistent or exceptional spelling-sound correspondences (e.g., HAVE, PINT). Thus, a clear extension of this experiment would be to manipulate the spelling-sound consistency as well as the frequency of the target words. Given that the derivation of phonology for low-frequency exception words is slowed relative to items higher in either frequency or consistency (Andrews, 1982; Seidenberg et al., 1984a; Taraban & McClelland, 1987; Waters & Seidenberg, 1985), we would expect to observe greater semantic priming when naming these items—that is, a three-way interaction of frequency, consistency, and priming context. This prediction mirrors the three-way interaction of frequency, consistency, and imageability in naming found by Strain et al. (1995). In both cases, the effect of semantics is greatest for items with the weakest spelling-sound mapping—low-frequency exception words.

A model of reading acquisition with differential development of the orthographic-phonological mapping versus the orthographic-semantic mapping would have important, but untested, developmental implications (also see Share, 1995). We have suggested that good readers show larger semantic priming effects in lexical decision than in naming (Keefe & Neely, 1990; Lorch et al., 1986; Lupker, 1984) because their well-developed spelling-sound mapping allows them to pronounce words rapidly, thereby reducing the effects of semantics on naming. However, poor readers should exhibit equal semantic priming effects in lexical decision and naming because their under-developed grapheme-phoneme connections allow semantic information to influence their slow naming processes. A differential development model also predicts that any factor which increases the ability of semantics to influence naming, such as reading low-frequency exception words at long SOAs, will the increase priming effects in naming for good readers. By contrast, these factors should minimally influence the magnitude of semantic priming effects in naming for poor readers because the orthographic system does not strongly drive the phonological system so semantics can influence naming regardless of frequency or SOA.

Notice that our argument here is very similar to the one formulated to explain the cross-linguistic semantic priming differences in naming. For example, Katz and Feldman (1983) and Frost, Katz, and Bentin (1987) compared semantic priming effects on naming in English with the effects in Serbo-Croatian, a more shallow orthography with greater spelling-sound consistency, and with the effects in unpointed Hebrew, a deeper, less consistent orthography. They found no semantic priming in Serbo-Croatian, and greater priming in Hebrew than in English. These findings can be understood as natural consequences of the basic properties of distributed network models, given their sensitivity to the relative degree of systematicity of the orthography-phonology mapping (see Seidenberg, 1992, for discussion).

Our empirical data and computational simulation showed that the adult condition exhibited about half as much priming as the child condition. Our finding of age-related decreases in the magnitude of semantic priming is supported by other empirical studies (Schwantes, 1981; Simpson & Lorsbach, 1983; West & Stanovich, 1978). It appears that, in English, higher-level (semantic) information influences word recognition less as children become skilled readers. By contrast, lower-level (orthographic and phonological) information appears to influence the reading process more as children develop. Only two published studies of single-word priming have directly examined the relative influences of orthographic and phonological processes in children's visual word recognition (Goswami, 1990; Hansen & Bowey, 1992), and these studies did not examine developmental differences. More recently, Booth et al. (submitted) have shown that there is a strong positive relationship between the magnitude of orthographic-phonological priming and both naming accuracy and age. Moreover, older and high-ability children can activate this orthographic and phonological information more quickly than younger and low-ability children. Unfortunately, no investigation to date has examined developmental differences in orthographic, phonological, and semantic priming effects in a single group of children.

One possibility is that beginning readers compensate for their deficient knowledge of spelling-sound correspondences by bringing to bear semantic knowledge about the world. However, as children learn the statistical regularities between phonology and orthography, they rely less on semantics and more on interactions between orthographic and phonological representations for rapid word recognition. These developmental differences have important implications for models of visual word recognition because reading acquisition does not consist simply of age-related increases in all component skills. Rather, some effects, such as semantic priming, appear to decrease with age in English, whereas other effects, such as orthographic and phonological priming, appear to increase with development.

Interestingly, Kang and Simpson (1996) have recently demonstrated a pattern of developmental effects in Korean, a shallow orthography like Serbo-Croatian, that are exactly the opposite of those found in English. Specifically, children learning to read Korean appear to exhibit a decrease in phonological priming and an increase in semantic priming with age. Korean-reading children may show greater phonological priming than English-reading children in the initial phases of learning to read because they benefit from the greater spelling-sound consistency of Korean. Over the course of development, however, mappings with semantics become stronger, so meaning has a larger influence on word recognition and this reduces the phonological priming effect. Unfortunately, to our knowledge, there is no well controlled cross-linguistic study that directly examines developmental differences in phonological and semantic priming.

Conclusion

Semantic priming phenomena have played a critical role in constraining theories of lexical processing. It is almost universally accepted that a comprehensive account of these phenomena must incorporate multiple mechanisms. Single-mechanism accounts of semantic priming phenomena were abandoned by most researchers because they did not seem capable of accounting for strategic effects (Neely, 1991). A central goal of the current work is to reconsider this conclusion in light of more recent progress in understanding the computational properties of distributed network models and in applying them to complex empirical phenomena.

Our empirical work demonstrates that frequency effects on semantic priming depend on perceptual ability. Our computational work demonstrates that this pattern of data is a natural consequence of the nonlinear effects within a distributed network model that derives the meanings of written words. The model also exhibits the shift from facilitation dominance at short SOA to inhibition dominance at long SOA, without recourse to expectancy-based processes. Moreover, other phenomena thought to implicate strategic processes may instead be explained in terms of shifts in response criteria for decision processes outside the lexical system. It is certainly the case that considerable work remains to be done in order to extend the current approach into a full account of semantic priming phenomena in lexical processing. Even so, the relative success of a single-mechanism, distributed network model in accounting for data which have heretofore been taken to necessitate additional, expectancy-based processes suggests that such models may provide a viable alternative to multiple-mechanism accounts of lexical processing.

Acknowledgments

The order of authorship is alphabetical. The research was supported by NIMH Training Grant 5T32MH19102, NICHD Grant 80258 to the first author, and an NIMH FIRST award (MH55628) to the second author. The computational simulation was run using customized software written within the Xerion simulator (version 3.1) developed by Drew van Camp, Tony Plate, and Geoff Hinton at the University of Toronto. We thank the students and administrators at Holton Arms School and Landon School for their participation in this study. We also thank Gregory Robison and Rwanda Latoya Jackson for their assistance in conducting the experiments, and Marlene Behrmann, Chris Kello, William Hall, Chuck Perfetti, Jay McClelland, Brian MacWhinney, and the CMU PDP research group for helpful comments and/or discussion. Correspondence regarding this article may be sent either to James Booth (jbooth+@andrew.cmu.edu), Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213–3890, USA, or to David Plaut (plaut@cmu.edu), Mellon Institute 115–CNBC, 4400 Forbes Avenue, Pittsburgh, PA 15213–2683, USA.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition*, *10*, 565–575.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 234–254.

Antos, S. J. (1979). Processing facilitation in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 527–545.

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583–612). New York: Academic Press.

Azuma, T., & Van Orden, G. C. (1997). Why SAFE is better than FAST: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language*, *36*, 484–504.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357.

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meanings in word recognition: A review of the literature. In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–222). Hillsdale, NJ: Erlbaum.

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 336–345.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 252–259.

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory and Cognition*, *8*, 493–512.

Becker, C. A. (1985). What do we really know about semantic context effects during reading? In D. Besner, T. G. Waller, & E. M. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (pp. 125–166). New York: Academic Press.

Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1059–1082.

Besner, D., & Joordens, S. (1995). Wrestling with ambiguity–Further reflections: Reply to Masson and Borowsky (1995) and Rueckl (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 515–301.

Besner, D., & Smith, M. C. (1992). Models of visual word recognition: When obscuring the stimulus yields a clearer view. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 468–482.

Biemiller, A. (1970). The development and use of graphic and contextual information as children learn to read. *Reading Research Quarterly*, *6*, 75–96.

Booth, J. R., Hunt, S. B., Perfetti, C. A., & MacWhinney, B. (1997). *The influence of rapid sequential perception on orthographic and phonological processing in reading disabled children and adults*. Paper accepted to Society for the Scientific Study of Reading, San Diego, CA.

Booth, J. R., Perfetti, C. A., & MacWhinney, B. (submitted). *Quick, automatic, and general activation of orthographic and phonological representations in young readers*. Manuscript submitted for publication.

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 813–840.

Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 63–85.

Chauvin, Y. (1988). *Symbol acquisition in humans and neural (PDP) networks*. PhD thesis, University of California, San Diego.

Chiarello, C., Burgess, C., Richards, L., & Pollock, A. (1990). Semantic and associative priming in the cerebral hemispheres: Some words do, some words don't...sometimes, some places. *Brain and Language*, *38*, 75–104.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory and Cognition*, *12*, 590–606.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*, 332–361.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

de Groot, A. M. B. (1984). Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target. *Quarterly Journal of Experimental Psychology*, *36A*, 253–280.

den Heyer, K. (1985). On the nature of the proportion effect in semantic priming. *Acta Psychologica*, *60*, 25–38.

den Heyer, K., Briand, K., & Dannenbring, G. (1985a). Automatic and strategic factors in semantic priming: An examination of Becker's model. *Memory and Cognition*, *13*, 228–232.

den Heyer, K., Briand, K., & Smith, L. (1985b). Automatic and strategic factors in semantic priming: An examination of Becker's model. *Memory and Cognition*, *13*, 228–232.

Detterman, D. K., & Daniel, M. H. (1991). Correlations of mental tests with each other and with cognitive variables are highest with low iq groups. *Intelligence*, *15*, 247–250.

Dosher, B. A., & Rosedale, G. (1989). Integrated retrieval cues as a mechanism for priming in retrieval from memory. *Journal of Experimental Psychology: General*, *2*, 191–211.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test—revised*. Circle Pines, MN: American Guidance Service.

Eden, G. F., Stein, J. F., Wood, H. M., & Wood, F. B. (1995). Temporal and spatial processing in reading disabled and normal children. *Cortex*, *31*, 451–468.

Eden, G. F., VanMeter, J. W., Rumsey, J. M., Maisog, J. M., Woods, R. P., & Zeffiro, T. A. (1996). Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature*, *382*, 66–69.

Epelboim, J., Booth, J. R., & Steinman, R. M. (1994). Reading unspaced text: Implications for theories of reading eye movements. *Vision Research*, *34*, 1735–1766.

Epelboim, J., Booth, J. R., & Steinman, R. M. (1996). Much ado about nothing: The place of space in text. *Vision Research*.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin and Review*, *2*, 460–493.

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory and Cognition*, *11*, 565–574.

Flores d'Arcais, G. B., Schreuder, R., & Glazenborg, G. (1985). Semantic activation during recognition of referential words. *Psychological Research*, *47*, 39–49.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 104–115.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recal. *Psychological Review*, *91*, 1–67.

Glanzer, M., & Ehrenreich, S. L. (1979). Structure and search of the internal lexicon. *Journal of Verbal Learning and Verbal Behaviour*, *18*, 381–398.

Goldsmith-Phillips, J. (1989). Word and context in reading development: A test of the interactive-compensatory hypothesis. *Journal of Educational Psychology*, *81*, 299–305.

Golinkoff, R. M., & Rosinki, R. R. (1976). Decoding, semantic processing, and reading comprehension skill. *Child Development*, *47*, 252–258.

Gordon, B. (1983). Lexical access and lexical decision: Mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behaviour*, *22*, 24–44.

Goswami, U. (1990). Phonological priming and orthographic analogies in reading. *Journal of Experimental Child Psychology*, *49*, 323–340.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*, 518–565.

Guttentag, R. E., & Haith, M. M. (1979). A developmental study of automatic word processing in a picture classification task. *Child Development*, *50*, 196–206.

Hansen, J., & Bowey, J. A. (1992). Orthographic rimes as functional units of reading in fourth-grade children. *Australian Journal of Psychology*, *44*, 37–44.

Henik, A., Freidrich, F. J., & Kellogg, W. A. (1983). The dependence of semantic relatedness effects upon prime processing. *Memory and Cognition*, *11*, 366–373.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1331–1356.

Hinton, G. E. (1989). Connectionist learning procedures. *Artificial Intelligence*, *40*, 185–234.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (pp. 77–109). Cambridge, MA: MIT Press.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Science, USA*, *79*, 2554–2558.

Jackson, N. E., & Biemiller, A. J. (1985). Letter, word, and text reading times of precocious and average readers. *Child Development*, *56*, 196–206.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *104*, 130–136.

Jared, D. (1997). Evidence that strategy effects in word naming reflect changes in output timing rather than changes in processing route. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1424–1438.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General*, *120*, 358–394.

Jastrzembski, J. (1981). Multiple mearnings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology*, *13*, 278–305.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin*, *96*, 29–44.

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1083–1105.

Joordens, S., & Besner, D. (1992). Priming effects that span an intervening unrelated word: Implications for models of memory representation and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 483–491.

Kang, H., & Simpson, G. B. (1996). Development of semantic and phonological priming in a shallow orthography. *Developmental Psychology*, *32*, 860–866.

Katz, W. F., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 157–166.

Kawamoto, A. (1988). Distributed representations of ambiguous words and their resolution in a connectionist network. In S. L. Small, G. W. Cottrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence*. San Mateo, CA: Morgan Kaufmann.

Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing approach. *Journal of Memory and Language*, *32*, 474–516.

Kawamoto, A. H., Farrar, W. T., & Kello, C. T. (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1233–1247.

Kawamoto, A. H., & Zemblige, J. H. (1992). Pronunciation of homographs. *Journal of Memory and Language*, *31*, 349–374.

Kaye, D. B., & Brown, S. W. (1985). Levels and speed of processing effects on word analysis. *Memory and Cognition*, *13*, 425–434.

Keefe, D. E., & Neely, J. H. (1990). Semantic priming in a pronunciation task: The role of prospective prime-generated expectancies. *Memory and Cognition*, *18*, 289–298.

Keil, F. C. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.

Keil, F. C. (1987). Conceptual development and category structure. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 175–200). Cambridge, U.K.: Cambridge University Press.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the time course of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 601–609.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lorch, R. F., Balota, D., & Stamm, E. (1986). Locus of inhibition effects in the priming of lexical decisions: Pre- or post-lexical access? *Memory and Cognition*, *14*, 95–103.

Lovegrove, W., Martin, F., & Slaghuis, W. (1986). A theoretical and experimental case for a visual deficit in specific reading disability. *Cognitive Neuropsychology*, *3*, 225–267.

Lukatela, G. (1994). Visual lexical access is initially phonological: II. Evidence from phonological priming by homophones and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 107–???

Lukatela, G., Savic, M., Urosević, Z., & Turvey, M. T. (1997). Phonological ambiguity impairs identity priming in naming and lexical decision. *Journal of Memory and Language*, *36*, 360–381.

Lukatela, G., & Turvey, M. T. (1990). Phonemic similarity effects and prelexical phonology. *Memory and Cognition*, *18*, 128–???

Lukatela, G., & Turvey, M. T. (1991). Phonological access of the lexicon: Evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 951–988.

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: I. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 107–128.

Lund, K., Burgess, C., & Atchley, R. A. (1995). Semantic and associative priming in high-dimensional semantic space. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 660–665). Hillsdale, NJ: Erlbaum.

Lupker, S. J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning and Verbal Behaviour*, *23*, 709–733.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 570–590.

Marchand, H. (Ed.). (1969). *The categories and types of present-day English word-formation*. Munich: C. H. Becksche Verlagsbuchhandlung.

Masson, M. E. J. (1991). A distributed memory model of context effects in word identfication. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 233–263). Hillsdale, NJ: Erlbaum.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 3–23.

Masson, M. E. J., & Borowsky, R. (1995). Unsettling questions about semantic ambiguity in connectionist models: Comment on Joordens and Besner (1994). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 509–514.

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 693–706.

McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. Cambridge, MA: MIT Press.

McClelland, J. L., Rumelhart, D. E., & the PDP Research Group (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models*. Cambridge, MA: MIT Press.

McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1155–1172.

McLeod, P., Plunkett, K., & Rolls, E. T. (in press). *Introduction to connectionist modelilng of cognitive processes*. Oxford, UK: Oxford University Press.

McNamara, T. P. (1992). Theories of priming I: Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1173–1190.

McNamara, T. P. (1994). Theories of priming II: Types of primes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 507–520.

McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language*, *27*, 545–559.

McRae, K., & Boisvert, S. (in press). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*, 99–130.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word recognition. In P. M. A. Rabbitt, & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press.

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory and Cognition*, *17*, 141–147.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148–197). Hillsdale, NJ: Erlbaum.

Monsell, S., Patterson, K., Graham, A., Hughes, C. H., & Milroy, R. (1992). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 452–467.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review*, *76*, 165–178.

Moss, H. E., Hare, M. L., Day, P., & Tyler, L. K. (1994). A distributed memory model of the associative boost in semantic priming. *Connection Science*, *6*, 413–427.

Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 863–883.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*, 609–626.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, Vol. 24 (pp. 207–248). New York: Academic Press.

Neely, J. H., Keefe, D. E., & Ross, K. (1989). Semantic priming in a lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1003–1019.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. (1994). *The University of South Florida word association, rhyme, and word fragment norms*. Unpublished manuscript.

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.

Paap, K. R., & Noel, R. W. (1991). Dual route models of print to sound: Still a good horse race. *Psychological Research*, *53*, 13–24.

Pearlmutter, B. A. (1989). Learning state space trajectories in recurrent neural networks. *Neural Computation*, *1*, 263–269.

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition*, *62*, 223–???

Perfetti, C. A., & Bell, L. (1991). Speech activation during the first 40 msec of word activation: Evidence from backward masking and masked priming. *Journal of Memory and Language*, *30*, 473–485.

Perfetti, C. A., Finger, E., & Hogaboam, T. (1978). Sources of vocalization latency differences between skilled and less skilled young readers. *Journal of Educational Psychology*, *70*, 730–739.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology*, *67*, 461–469.

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. Just, & P. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141–183). Hillsdale, NJ: Erlbaum.

Peterson, R. R., & Simpson, G. B. (1989). Effect of backward priming on word recognition in single-word and sentence contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1020–1032.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 37–42). Hillsdale, NJ: Erlbaum.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of naming and lexical decision. *Language and Cognitive Processes*, *12*.

Plaut, D. C., & Booth, J. R. (in preparation). *Individual and developmental differences in semantic priming of word naming: Empirical evidence and a computational model*. Manuscript in preparation.

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

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt, & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press.

Postman, L., & Keppel, G. (1970). *Norms of word associations*. New York: Academic Press.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science*, *12*, 410–430.

Quinlan, P. (1991). *Connectionism and psychology: A psychological perspective on new connectionist research*. Chicago: University of Chicago Press.

Raduege, T. A., & Schwantes, F. M. (1987). Effects of rapid word recognition training on sentence context effects in children. *Journal of Reading Behavior*, *19*, 395–414.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *88*, 385–408.

Rueckl, J. G. (1995). Ambiguity and connectionist networks: Still settling into a solution–Comment on Joordens and Besner (1994). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 501–508.

Rumelhart, D. E., Durbin, R., Golden, R., & Chauvin, Y. (1995). Backpropagation: The basic theory. In Y. Chauvin, & D. E. Rumelhart (Eds.), *Back-propagation: Theory, architectures, and applications* (pp. 1–34). Hillsdale, NJ: Erlbaum.

Schadler, M., & Thissen, D. M. (1981). The development of automatic word recognition and reading skill. *Memory and Cognition*, *9*, 132–141.

Schwanenflugel, P. J., & Rey, M. (1986). Interlingual semantic facilitation: Evidence for a common representational system in the bilingual lexicon. *Journal of Memory and Language*, *25*, 605–618.

Schwantes, F. M. (1981). Locus of the context effect in children's word recognition. *Child Development*, *52*, 895–903.

Schwantes, F. M. (1985). Expectancy, integration, and interactional processes: Age differences in the nature of words affected by sentence context. *Journal of Experimental Child Psychology*, *39*, 212–229.

Schwantes, F. M. (1991). Children's use of semantic and syntactic information for word recognition and determination of sentence meaningfulness. *Journal of Reading Behavior*, *23*, 335–350.

Schwantes, F. M., Boesl, S. L., & Ritz, E. G. (1980). Children's use of context in word recognition: A psycholinguistic guessing game. *Child Development*, *51*, 730–736.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 876–900.

Seidenberg, M. S. (1992). Beyond orthographic depth: Equitable division of labor. In R. Frost, & K. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 85–118). Amsterdam: Elsevier.

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

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984a). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behaviour*, *23*, 383–404.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984b). Pre- and postlexical loci of contextual effects on word recognition. *Memory and Cognition*, *12*, 315–328.

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*, 151–218.

Sharkey, A. J., & Sharkey, N. E. (1992). Weak contextual constraints in text and word priming. *Journal of Memory and Language*, *31*, 543–572.

Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1191–1210.

Simpson, G. B., & Lorsbach, T. C. (1983). The development of automatic and conscious components of contextual facilitation. *Child Development*, *54*, 760–772.

Simpson, G. B., & Lorsbach, T. C. (1987). Automatic and conscious context effects in average and advanced readers. *Journal of Research in Reading*, *10*, 102–112.

Smith, L. C., Briand, K., Klein, R. M., & den Heyer, K. (1987). On the generality of Becker's Verification Model. *Canadian Journal of Experimental Psychology*, *41*, 379–386.

Smith, M. C., Besner, D., & Miyoshi, H. (1994). New limits to automaticity: Context modulates semantic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 104–115.

Smith, M. C., Theodor, L., & Franklin, P. E. (1983). On the relationship between contextual faciliation and depth of processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 697–712.

Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (pp. 194–281). Cambridge, MA: MIT Press.

Stahl, S. A., Hare, V. C., Sinatra, R., & Gregory, J. F. (1991). Defining the role of prior knowledge and vocabulary in reading comprehension. *Journal of Reading Behavior*, *23*, 487–508.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*, 32–71.

Stanovich, K. E., Nathan, R. G., West, R. F., & Vala-Rossi, M. (1985). Children's word recognition in context: Spreading activation, expectancy, and modularity. *Child Development*, *56*, 1418–1428.

Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 658–672.

Stanovich, K. E., West, R. F., & Feeman, D. (1981). A longitudinal study of sentence context effects in second-grade children: Tests of an interactive-compensatory model. *Journal of Experimental Child Psychology*, *32*, 185–199.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologia*, *30*, 276–315.

Stolz, J. A., & Besner, D. (1997). Visual word recognition: Effort after meaning but not (necessarily) meaning after effort. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1314–1322.

Stolz, J. A., & Neely, J. H. (1995). When target degradation does and does not enhance semantic context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 596–611.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 744–774.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language*, *36*, 337–359.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140–1154.

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. In P. Tallal, A. M. Galaburda, R. R. Llinas, & E. von Euler (Eds.), *Temporal information processing in the nervous system: Special reference to dyslexia and dysphasia.* (pp. 27–47). New York: New York Academy of Sciences.

Tallal, P., & Piercy, M. (1973). Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, *11*, 389–398.

Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*, *12*, 83–93.

Tallal, P., & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia*, *13*, 69–74.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word recognition. *Journal of Memory and Language*, *26*, 608–631.

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 and Cognition*, *5*, 84–99.

Usher, M., & McClelland, J. L. (1995). *On the time course of perceptual choice: A model based on principles of neural computation* (Technical Report PDP.CNS.95.5). Pittsburgh, PA: Carnegie Mellon University, Department of Psychology.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound and reading. *Memory and Cognition*, *15*, 181–198.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1269.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from the spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 371–386.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*, 488–522.

Vernon, P. A. (1987). *Speed of information processing and intelligence*. Norwood, NJ: Ablex.

Verstaen, A., Humphreys, G. W., Olson, A., & d'Ydewalle, G. (1995). Are phonemic effects in backward masking evidence for automatic prelexical phonemic activation in visual word recognition? *Journal of Memory and Language*, *34*, 335–356.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time course and decision criteria. *Memory and Cognition*, *13*, 557–572.

Wechsler, D. (1991). *Wechsler intelligence scale for children—Third edition*. San Antonio, TX: Harcourt, Brace & Company.

West, R. F., & Stanovich, K. E. (1978). Automatic contextual facilitation in readers of three ages. *Child Development*, *49*, 717–727.

Wijnen, F. (1992). Incidental word and sound errors in young speakers. *Journal of Memory and Language*, *31*, 734–755.

Appendix 1: Nonword Primes, Unrelated Primes, Related Primes, and Target Words Used in Experiments 1–3

Nonword Primes	Unrelated Primes	Related Primes	Target Words	Nonword Primes	Unrelated Primes	Related Primes	Target Words
KARBS	EIGHT	ABOVE	BELOW	YAMBE	FINAL	ADULT	CHILD
DOFER	CRAWL	AGONY	PAINS	KARPE	SCREW	ALARM	CLOCK
RAIFY	SPLIT	ARGUE	FIGHT	STELT	BASIC	BEING	HUMAN[d]
TREGS	ALONE	BIRTH	DEATH[d]	FWIST	GUARD	BLADE	KNIFE[d]
VIGHT	VENUS	BLANK	EMPTY	SLELS	FAVOR	BLAZE	FIRES
STELI	TRUNK	BORED	TIRED[d]	DILCH	CRACK	BRIDE	GROOM
MEASH	LOWER	BRIEF	SHORT	EROWN	EARLY	BRING	TAKES
CHESA	TOWER	CANOE	BOATS	SHILT	NERVE	CHAIN	LINKS
SLOVE	MOIST	CHUCK	THROW[d]	JAMOR	GIANT	CIGAR	SMOKE
KOUGH	UNITE	CLEAN	DIRTY[d]	DIGRI	BEGIN	CLOSE	OPENS
RIFEY	SHOOT	COACH	TEAMS[d]	REVRI	ARROW[c]	CORAL	REEFS
SUMIC	CROWD	COURT	JUDGE	GLANE	TOPIC	CRANE	LIFTS[d]
PEESH	USUAL	CREEK	RIVER	DORCH	COUNT	CYCLE	BIKES
SHALS	CHEST	DEATH	LIVES	SLIND	STRAW	DITCH[b]	HOLES[d]
KESPO	PLAIN	DONOR	BLOOD[d]	SMONT	BLIND	ENTER	EXITS[d]
SHEDA	BENCH[c]	FAIRY	TALES	OMOSE	ALIKE[c]	FENCE[b]	POSTS
YESRA	CREAM	FLAME	FIRES[d]	OCKSO	TODAY	FLOOD	WATER
ROBAD	PATCH	FRESH	FRUIT[d]	TOLBS	HURRY	FUNNY	LAUGH[d]
WRUPP	CLIMB	GHOUL	GHOST	VANGE	SIGHT	GLOVE	HANDS
LAVUE	WIDTH	GRAIN	WHEAT	THEET	SCORE	GRASP	HOLDS
SKALT	EVENT	GRASS	GREEN	BISER	READY	HEAVY	LIGHT[d]
STEAF	PRIZE	HONEY	SWEET	DRAFU	SHAPE	HOUSE	HOMES
DUTSY	ALLOW	JOINT	KNEES	SALGS	LEVEL	KNOCK	DOORS[a]
TORMS	NEVER	LABOR	WORKS	EKAPS	PARTY	LARGE	SMALL[a]
TRUIF	ROUGH	LEMON	LIMES[d]	FIETH	ANGLE	LOOSE[b]	TIGHT
WEASH	SHINE	MAJOR	MINOR	SYAMP	BEAST	MAPLE	TREES
EOUSH	PITCH	MARCH	APRIL	SNOGS	STEAM	MINTS	CANDY[d]
POUGH	COLOR	MONTH	YEARS[d]	CARCK	CHECK	MOTEL	HOTEL
SLAKE	CAUSE	NORTH	SOUTH	GOWAN	CHEEK	NOVEL	BOOKS
JUDIT	SOLID	PAINT	BRUSH	KLOPS	CABIN	PASTE	GLUES
APULT	FAITH[c]	PAUSE	STOPS	KABES	EXTRA	PHONE	CALLS
RUESH	DENSE[c]	PHONY	FAKES[a]	GLAFS	REPLY	PIANO	PLAYS
BLOGE	HABIT	PILOT	PLANE	SPLAY	STALK	POKER[b]	CARDS[d]
SHOAT	LEAVE	PRINT	WRITE	PRAPE	PEARL	QUACK	DUCKS[d]
NAIRT	SHOCK	QUEEN	KINGS[d]	KULLS	CLEAR	RADIO	MUSIC
FOROL	CLOTH	RAZOR	SHARP	SCRIE	WORSE	REACH	GRABS[a]
NOACE	STIFFC	SCENT	SMELL	RENGE	SWIFT	SHAME	GUILT
RAPEL	VOICE	SHARE	GIVES	GWINS	GOING	SHEET	PAPER
THRON	TENSE[c]	SHIFT[b]	GEARS[a]	LANEG	PUPIL	SHIRT	PANTS
OPINT	NOTES	SHORE	BEACH	GANRY	BURST	SHOUT	YELLS
MILTI	DRINK	SKIRT	DRESS	HASLY	AHEAD	SLICE	PIECE[d]
HIFTA	CHARM	SMILE	HAPPY	GATCH	PROUD	SNAKE	BITES[a]
CHIRD	RAISE	SOCKS	SHOES	LUPPE	DREAM	SOUND	NOISE
WETCH	STORE	SPARE	TIRES	MOTTU	AVOID	SPEAK	TALKS
ROWEL	FLOOR	SPEND	MONEY	HESET	RAPID	SPOON	FORKS
LINDS	QUICK	STALL	HORSE	NACLE	ANGER	STARE	LOOKS[d]
VOBAE	MOTOR	STEEL[b]	METAL	TILOP	CURVE	STILL	MOVES
CANFY	STAMP	STONE	ROCKS	RAICH	BOUND	STORM	RAINS

continued on next page

BLACE	STAND	STUFF	THING	STANT	STATE	SUPER	GREAT
NUIST	CRASH	SWEAR	CURSE[a]	SNISP	DRILL	SWEEP	BROOM[a]
TOOFA	RIFLE	TABLE	CHAIR	ODEAS	ADMIT	TEACH	LEARN[d]
LEJLY	GUEST	THIEF	STEAL	FIRCH	METER	TIGER	LIONS
RUSOI	FROST	TOAST	BREAD	GAMIK	GLORY	TOOTH	DECAY
DREAB	VISIT	TOUCH	FEELS[d]	GELEA	NURSE	TRAIL	PATHS
ECHAT	SCALE	TRAIN	TRACK[d]	ROCAL	EQUAL	TRICK	TREAT
TAFAL	TOTAL	TRUCE	PEACE[d]	SNOLE	MODEL	TWIST	TURNS
VIGES	APART[c]	UNCLE	AUNTS[a]	JIETZ	SWAMP	WAGON	WHEEL[d]
CHUTH	PLATE	WAVES[b]	OCEAN	SMOUL	CHIEF	WHITE	BLACK[d]
DUKOL	CHINA	WINGS	BIRDS[d]	QUARF	CLOUD	WRIST	WATCH
LERRI	FOUND	WRONG	RIGHT	KLIGS	FRONT	YOUTH	YOUNG[d]

Note. All items that were outliers or yielded no reliable priming effect in Experiment 1 with college students were eliminated from the word list for Experiment 2 with the elementary students.

[a] Outlier in nonword condition.

[b] Outlier in related condition.

[c] Outlier in unrelated condition.

[d] No reliable priming effect.

Appendix 2: Mean Reaction Times (and Standard Deviations) for Subjects from Experiments 1–3 and for the Network

Subject Group	High-Frequency Targets			Low-Frequency Targets		
	Related	Unrelated	Nonword	Related	Unrelated	Nonword
Adults						
Short SOA						
High Perceptual Ability						
Subjects	601 (56)	618 (61)	643 (65)	613 (63)	665 (78)	692 (76)
Network	650 (40)	669 (35)	675 (40)	690 (35)	717 (29)	723 (42)
Low Perceptual Ability						
Subjects	678 (69)	705 (90)	742 (78)	713 (94)	735 (74)	744 (73)
Network	656 (63)	686 (68)	681 (78)	729 (80)	754 (69)	728 (68)
Long SOA						
High Perceptual Ability						
Subjects	686 (54)	693 (48)	692 (44)	694 (48)	727 (61)	720 (56)
Network	696 (36)	711 (31)	708 (33)	731 (33)	762 (27)	757 (30)
Low Perceptual Ability						
Subjects	760 (61)	780 (64)	756 (56)	789 (68)	806 (70)	796 (65)
Network	712 (61)	766 (106)	729 (80)	814 (129)	843 (91)	801 (76)
Children						
Short SOA[a]						
High Perceptual Ability						
Network	721 (56)	743 (52)	750 (55)	822 (100)	866 (114)	867 (108)
Low Perceptual Ability						
Network	742 (105)	791 (133)	775 (115)	968 (221)	1007 (217)	965 (200)
Long SOA						
High Perceptual Ability						
Subjects	837 (95)	855 (100)	868 (87)	861 (95)	913 (97)	920 (92)
Network	779 (47)	792 (45)	771 (51)	878 (101)	940 (125)	892 (103)
Low Perceptual Ability						
Subjects	956 (83)	1004 (91)	992 (79)	1014 (88)	1059 (137)	1064 (129)
Network	843 (120)	927 (168)	826 (133)	1062 (205)	1142 (196)	1010 (192)

[a] Note that the empirical studies did not test children at the short SOA.