Psychological Review
2012, Vol. 119, No. 2, 417–430

0033-295X/12/$12.00 DOI: 10.1037/a0027175

THEORETICAL NOTE

Competition and Cooperation Among Similar Representations: Toward a Unified Account of Facilitative and Inhibitory Effects of Lexical Neighbors

Qi Chen and Daniel Mirman
Moss Rehabilitation Research Institute, Elkins Park, Pennsylvania

One of the core principles of how the mind works is the graded, parallel activation of multiple related or similar representations. Parallel activation of multiple representations has been particularly important in the development of theories and models of language processing, where coactivated representations (*neighbors*) have been shown to exhibit both facilitative and inhibitory effects on word recognition and production. Researchers generally ascribe these effects to interactive activation and competition, but there is no unified explanation for why the effects are facilitative in some cases and inhibitory in others. We present a series of simulations of a simple domain-general interactive activation and competition model that is broadly consistent with more specialized domain-specific models of lexical processing. The results showed that interactive activation and competition can indeed account for the complex pattern of reversals. Critically, the simulations revealed a core computational principle that determines whether neighbor effects are facilitative or inhibitory: Strongly active neighbors exert a net inhibitory effect, and weakly active neighbors exert a net facilitative effect.

Keywords: interactive activation, competition, neighbor effects, neighborhood effects, lexical processing

Supplemental materials: http://dx.doi.org/10.1037/a0027175.supp

Across theoretical frameworks and domains of cognition, one of the core principles of how the mind works is the graded, parallel activation of multiple related or similar representations. This idea is a central tenet of parallel distributed processing models of cognition (e.g., McClelland, 1993; Rumelhart, McClelland, & the Parallel Distributed Processing Research Group, 1986), exemplar models of memory (e.g., Kalish, Lewandowski, & Kruschke, 2004; Medin & Schaffer, 1978), and Bayesian approaches to cognition (e.g., Griffiths, Kemp, & Tenenbaum, 2008). Parallel activation of multiple representations has been particularly important in the development of theories and models of language processing, where these related representations have been called *neighbors* and studied extensively in a wide range of tasks and contexts. Although there is broad agreement on the principle of parallel activation, the consequences are quite varied, and to date, there is no formal unified account of why lexical neighbors exert facilitative effects in some contexts and inhibitory effects in others. This is because research on neighbor effects has been almost entirely isolated by domain: There are sophisticated models of, for example, reading aloud (e.g., the dual route cascaded model: Coltheart, Davelaar, Jonasson, & Besner, 2001; the connectionist dual process model: Perry, Ziegler, & Zorzi, 2007; and the self-organizing lexical acquisition and recognition (SOLAR) model: Davis, 2010), which capture neighbor effects in reading aloud in great detail but do not address neighbor effects in picture naming, spoken word recognition, or other tasks. Here, we take an orthogonal approach: Instead of building a detailed model of one task, we build a simple and general model designed to be applicable across domains and use it uncover the computational principle that underlies the contrasting results in the literature. A corollary of this approach is that our model is not meant to compete with existing models of lexical processing or neighborhood effects; rather, it is meant to be a step toward bridging across existing models by identifying underlying computational principles that would need to hold in a full cross-domain model of lexical processing.

We first briefly review the literature on lexical neighbor effects, documenting which contexts elicit facilitative effects and which contexts elicit inhibitory effects. We then describe a simple interactive activation and competition (IAC) framework (e.g., McClelland & Rumelhart, 1981) for exploring these effects. Many researchers have used IAC frameworks to explain lexical neighborhood effects, either by intuition or through implemented computational models, but there has been little effort to explain why the same framework would predict inhibitory neighbor effects in some cases and facilitative neighbor

This article was published Online First February 20, 2012.

Qi Chen and Daniel Mirman, Moss Rehabilitation Research Institute, Elkins Park, Pennsylvania.

This research was supported by Grant R01DC010805 to Daniel Mirman and by the Moss Rehabilitation Research Institute. We thank Gary Dell for helpful comments on an early draft of this article. Full model code is available in the supplemental materials (http://dx.doi.org/10.1037/a0027175.supp).

Correspondence concerning this article should be addressed to Qi Chen, Moss Rehabilitation Research Institute, 50 Township Line Road, Elkins Park, PA 19027. E-mail: chenqi.research@gmail.com

effects in other cases. Using simulations, we show that these contrasts can arise in a single computational framework and provide a unified account of why some neighbors facilitate and others inhibit word processing.

Visual Word Recognition

Facilitative Effects of Orthographic and Phonological Neighbors

One of the first demonstrations of neighborhood effects in language processing was the finding that printed words with many orthographic neighbors were recognized more quickly than words with few neighbors (Andrews, 1989, 1992; Forster & Shen, 1996; Johnson & Pugh, 1994; Sears, Hino, & Lupker, 1995; for a review, see Andrews, 1997). Most of these studies used a lexical decision or naming (reading aloud) task, though the latter task also involves word production processes that, as we review later, are also sensitive to neighborhood effects. The typical definition of an orthographic neighborhood is Coltheart's *N* (Coltheart et al., 1977), which is the number of words that can be created by changing a single letter of a target word. For example, the target word *mint* has neighbors such as *mitt*, *tint*, *lint*, and *pint*. Other researchers have proposed alternative measures of orthographic neighborhood (e.g., Yarkoni, Balota, & Yap, 2008), but all of the measures have retained the core notion of letter-based similarity and have demonstrated facilitative effects of orthographic neighbors.[1]

Phonological neighbors—words than can be created by changing a single phoneme—can be somewhat different from orthographic neighbors. For example, the phonological neighborhood of *mint* includes orthographic neighbors, such as *tint* and *lint*, as well as additional words, such as *mince*, *meant*, and *minnow*, and excludes some orthographic neighbors, such as *pint*. Nevertheless, phonological neighbors also exert facilitative effects on visual word recognition (Yates, 2005; Yates, Locker, & Simpson, 2004).

Inhibitory Effects of Higher Frequency Neighbors

The studies just reviewed suggest cooperation among lexical neighbors during visual word recognition. Because word frequency facilitates word recognition, one might imagine that neighbors that are more frequent would be more facilitative. The data reveal the opposite pattern: When neighbors are more frequent than the target word, they exert an inhibitory effect on target word recognition (Davis, Perea, & Acha, 2009; Ferraro & Hansen, 2002; Grainger, 1990; Grainger & Jacobs, 1996; Grainger, O'Regan, Jacobs, & Segui, 1989, 1992; Grainger & Segui, 1990).

These results indicate that there must be some balance between facilitative and inhibitory effects of orthographic neighbors. The need for this kind of balance is further demonstrated by inhibitory effects of transposed letter neighbors: Words that have a transposed letter neighbor, such as *salt–slat*, are recognized more slowly than matched words that do not have a transposed letter neighbor, such as *halt* (Acha & Perea, 2008; Andrews, 1996; Johnson, 2009). Andrews (1996) showed that both the facilitative effect of letter substitution neighbors and the inhibitory effect of transposed letter neighbors can be demonstrated within a single data set and concluded, "It remains to be seen whether it is possible to find a single set of parameters that allow successful simulation of both phenomena" (p. 795). The search for models and parameters that can account for these contrasting patterns has been a major theme in the development of theories and models of visual word processing (for a review, see Grainger, 2008).

Spoken Word Recognition: Inhibitory Effects of Neighbors

When words are presented auditorily, rather than visually, the effect of neighbors reverses to inhibit word recognition (e.g., Luce, 1986; Luce & Pisoni, 1998). Indeed, neighborhood probability—a measure that combines relative word frequency and number of neighbors—accounts for about 15% of the variance in tasks like lexical decision and word repetition (Luce, 1986; Luce & Pisoni, 1998). The next best predictor is frequency alone, which only accounts for about 5% of the variance. The inhibitory effect of phonological neighbors on spoken word recognition has also been shown in word-to-picture matching (e.g., Magnuson, Dixon, Tanenhaus, & Aslin, 2007), gating (e.g., Garlock, Walley, & Metsala, 2001), priming (Dufour & Peereman, 2003a, 2003b; Goldinger, Luce, & Pisoni, 1989), and other paradigms. The effect has also been shown using different definitions of phonological neighbors (e.g., Luce & Pisoni, 1998; Magnuson et al., 2007; see also Benkí, 2003).

Although most evidence shows inhibitory effects of lexical neighbors on spoken word recognition, partial activation of many lexical representations may have facilitative effects at sublexical levels of speech processing. The best evidence of such facilitative top-down effects is a bias to identify ambiguous phonemes toward a denser lexical neighborhood (Newman, Sawusch, & Luce, 1997; see also Boyczuk & Baum, 1999). For example, an ambiguous sound between /g/ and /k/ was more likely to be identified as /g/ when followed by *–ice*, presumably because *gice* has higher neighborhood density than *kice*. Sublexical facilitation has also been implicated in facilitative effects of phonotactic probability—the relative likelihood of phoneme pairs (e.g., Vitevitch & Luce, 1998, 1999; Vitevitch, Luce, Pisoni, & Auer, 1999; see also Luce & Large, 2001). However, other researchers have questioned this result on the basis of failures to replicate (Lipinski & Gupta, 2005; T. Strauss, personal communication, May 5, 2009; see also Roodenrys & Hinton, 2002).

Spoken Word Production: Facilitative Effects of Neighbors

Considering the opposite effects of lexical neighbors in visual and spoken word recognition, one might imagine that the critical difference is the modality itself—that something about spoken language makes neighbor effects inhibitory. This hypothesis is inconsistent with the robust facilitative effects of lexical neighbors on spoken word production (e.g., Gordon, 2002; Kittredge, Dell, Verkuilen, &

[1] The facilitative effects of neighbors on visual word recognition appear to be stronger for low-frequency words than for high-frequency words (e.g., Coltheart et al., 2001; Yarkoni et al., 2008; see also Davis, 2010, who made the even stronger claim that there are no effects of neighbors on high-frequency words). Like the well-known Frequency × Regularity interaction, this could arise simply because recognition of high-frequency words is fast, making it is difficult to detect a neighbor facilitation effect (for a related discussion, see, e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; and for a similar account in a very different domain, see the discussion of the asymmetry of word and ink color interference effects in the Stroop task in Cohen, Dunbar, & McClelland, 1990).

Schwartz, 2008; Middleton & Schwartz, 2010; Vitevitch, 1997, 2002; Vitevitch & Sommers, 2003). These facilitative effects have been shown in error rates from natural speech corpora (e.g., Vitevitch, 1997), error elicitation tasks (e.g., Vitevitch, 2002), and picture naming tasks with healthy control participants (Mirman, Kittredge, & Dell, 2010; Vitevitch & Sommers, 2003). The effects have also been shown with aphasic speakers (e.g., Gordon, 2002; Kittredge et al., 2008; Middleton & Schwartz, 2010; Mirman et al., 2010). Similar patterns have been observed in response times (e.g., Vitevitch & Sommers, 2003), and dense lexical neighborhoods seem to help in avoiding tip-of-the-tongue states (i.e., words from sparse neighborhoods are more likely to cause tip-of-the-tongue states; e.g., Harley & Brown, 1998; Vitevitch & Sommers, 2003).

As discussed earlier, phonological neighbors appear to exert opposite effects in visual and spoken word recognition. The same pattern does not appear to hold for word production: Analyses of spoken and written spelling by aphasic participants showed that neighbors facilitated the successful production of a target word in either modality (Goldrick, Folk, & Rapp, 2010).

Facilitative effects of lexical neighbors in reading aloud could be due to facilitation at the visual word recognition level or the spoken word production level. However, researchers have argued that this effect is driven by the (phonological) word production aspect of the task rather than the (orthographic) word recognition aspect (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Mulatti, Reynolds, & Besner, 2006).

sDell and Gordon (2003) described a preliminary computational account of why lexical neighbors facilitate spoken word production but inhibit spoken word recognition. Their account, which was implemented and tested in simulations of the two-step interactive-activation model of lexical access (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997), was based on two key insights. The first insight was that phonological neighbor effects in spoken word production require interactivity, that is, bidirectional excitation between different levels, specifically between lexical and phonological representations. For phonologically similar facilitative *gangs* (cf. McClelland & Rumelhart, 1981; Taraban & McClelland, 1987) of words to become active, phoneme activation must feed back to lexical levels to activate them. The second insight was that word production is a semantically driven task and word recognition is a phonologically driven task. As a result, the strongest lexical competitors during word production are semantic neighbors, not phonological neighbors. Therefore, the (weak) activation of some phonological neighbors should not substantially increase ambiguity. The notion that phonological neighbors facilitate word production by helping the target word overcome competition from semantic neighbors is also supported by reduced semantic and omission error proportions for high phonological neighborhood density words (Kittredge et al., 2008; see also Middleton & Schwartz, 2010). In contrast, phonological neighbors are the strongest competitors in spoken word recognition; therefore, having more of them (i.e., a dense neighborhood) should substantially increase ambiguity and slow down recognition. As we show, this distinction between strong and weak competitors is at the core of determining whether neighbors facilitate or inhibit processing.

Semantic Neighbors: Opposite Effects of Near and Distant Neighbors

Following the same logic as defining neighbors in terms of form similarity (phonological or orthographic), neighbors can also be defined in terms of meaning (semantic) similarity. Several studies have found that words with many semantic neighbors or denser semantic neighborhoods are recognized more quickly (Buchanan, Westbury, & Burgess, 2001; Duñabeitia, Avilés, & Carreiras, 2008; Locker, Simpson, & Yates, 2003; Siakaluk, Buchanan, & Westbury, 2003; Yates, Locker, & Simpson, 2003). A finer grained analysis of the effects of semantic neighbors on visual word recognition (concreteness judgment task) found that distant semantic neighbors—concepts that share a few semantic features—facilitated word recognition. In contrast, near semantic neighbors—concepts that share many semantic features—inhibited word recognition (Mirman & Magnuson, 2008). A subsequent study replicated this finding in a word production (picture naming) task testing aphasic and speeded control participants (Mirman, 2011).

Although these opposite effects of near and distant semantic neighbors have not been investigated as thoroughly as other neighbor effects, they provide important constraining evidence. Reversals across the other domains could possibly be attributed to differences in domain-specific representations (e.g., orthography vs. phonology) or tasks (e.g., word recognition vs. production). The semantic neighbor effects tell a different story: The neighbors are all of the same type (i.e., semantic), and their effects are consistent across tasks (both word recognition and production), but near semantic neighbors exert inhibitory effects, whereas distant semantic neighbors exert facilitative effects.

The Present Study

Table 1 provides a summary of the qualitative effects of different kinds of lexical neighbors in different tasks, as reviewed

Table 1
Effect of Different Kinds of Neighbors in Different Tasks

Neighbor/task type	Behavior	Model
Form neighbors		
Visual word recognition	Facilitation	Facilitation (Figure 3, left)
Visual word recognition with high-frequency neighbor	Inhibition	Inhibition (Figure 3, middle)
Spoken word recognition	Inhibition	Inhibition (Figure 3, right)
Spoken word production	Facilitation	Facilitation (Figure 4)
Semantic neighbors		
Near neighbors	Inhibition	Inhibition (Figure 5, top)
Distant neighbors	Facilitation	Facilitation (Figure 5, bottom)

earlier. Although greatly simplified, this summary captures the fact that neighborhood effects are considered among the most robust findings in each domain notwithstanding that, across domains, these robust effects go in opposite directions. This isolation of domains has arisen because most researchers have focused on only one kind of neighbor effect or context (e.g., spoken word recognition and not visual word recognition or spoken word production). However, they have almost universally appealed to IAC to account for their findings. Therefore, a unified account may be possible. This is the goal of the present study: to empirically investigate the dynamics of IAC in different word-processing contexts to (a) examine whether the basic principles of IAC can correctly predict the direction of lexical neighbor effects and (b) uncover how or why the task and neighbor type may determine the direction of the effect.

Because our focus was specifically on the dynamics of IAC, we used a very simple implementation of IAC principles to maximize the tractability of the simulations. Consequently, because of the minimal nature of the model and the importance of the qualitative reversals in the behavioral data, we focus on the qualitative patterns produced by the model (see also Pitt, Kim, Navarro, & Myung, 2006, for a discussion of the value of global qualitative evaluation of computational models).

Simulations

Network Architecture

Although the simple IAC model described here was not intended to be a full model of lexical processing, IAC is a core principle of parallel distributed processing models in general, and as a result, the model architecture is closely related to leading models in each of the relevant domains. Specifically, it is closely related to models of visual word recognition (e.g., Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981), spoken word recognition (e.g., Gaskell & Marslen-Wilson, 1997; Magnuson, Tanenhaus, Aslin, & Dahan, 2003; McClelland & Elman, 1986), spoken word production (e.g., Dell et al., 1997), and lexical semantics (e.g., Cree, McRae, & McNorgan, 1999; Rogers & McClelland, 2004). Our modeling approach specifically builds on recent efforts to use simple IAC models to account for opposite effects of lexical neighbors on spoken word production versus recognition (Dell & Gordon, 2003) and on spoken versus visual word recognition (Magnuson & Mirman, 2007).

The basic structure of the model is shown in Figure 1. The model consisted of simple processing units organized into three layers: units in the first layer corresponded to elements of word form (i.e., phonemes or letters), units in the second layer corresponded to lexical elements (i.e., words in the model's lexicon), and units in the third layer corresponded to elements of meaning (i.e., semantic features of concepts denoted by the words). As in other IAC models (e.g., Dell et al., 1997; McClelland & Elman, 1986; McClelland & Rumelhart, 1981), congruent units in different layers were connected by bidirectional weighted connection links. That is, each word unit had bidirectional connections to its constituent letters or phonemes and to its semantic features.

To implement competition, units in the word layer were connected by bidirectional inhibitory connections. The inhibitory connection strength was scaled by a sigmoid function of unit activation, as shown in Figure 2. In other words, weakly active word units had very little inhibitory effect on other word units, and strongly active words units had a very strong inhibitory effect on other word units. This nonlinearity was implemented to allow initial parallel activation of many word candidates while still forcing the model to eventually settle to a single active representation. A similar approach was implemented by Cisek (2006) in a neurally based computational model of action planning and selection. In essence, this approach implements a decision-making or response selection mechanism similar to the Luce (1959) choice rule (see also Mirman, Yee, Blumstein, & Magnuson, 2011, for a sigmoid representation of the Luce choice rule) and that captures the notion of progressively increasing pressure to settle to a single active candidate (for related discussion of progressively increasing decision pressure, see Allopenna, Magnuson, & Tanenhaus, 1998; McRae, Spivey-Knowlton, & Tanenhaus, 1998).

To mimic the kind of recurrent connections that would be learned by semantic representations (e.g., Cree et al., 1999; Rogers & McClelland, 2004), semantic units generally had inhibitory connections, but this inhibition was reduced for each concept in

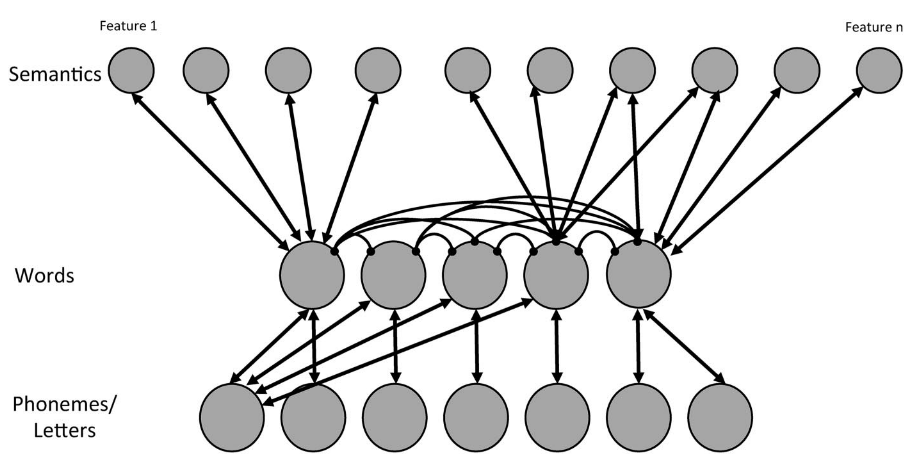

Figure 1. The full structure for the model.

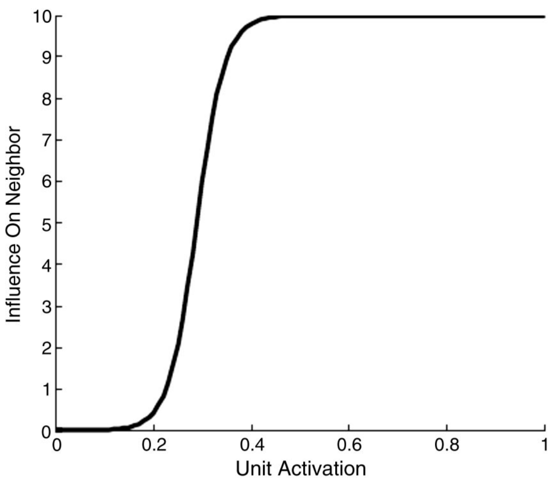

Figure 2. Sigmoid inhibition strength function.

which the semantic units (features) co-occurred (for evidence of facilitative effects of feature co-occurrence, see, e.g., Cree & McRae, 2003; Rogers & McClelland, 2004). In other words, a semantic feature such as “has wings” was assumed to have inhibitory connections to unrelated features, such as “has strings,” excitatory connections to strongly (cor)related features, such as “has feathers,” and intermediate weights to weakly related features, such as “made of metal” (airplane). Implementation details of these within-level connections at the semantic level are described in Simulation 5.[2] All units followed the standard IAC activation function in which positive net input drives unit activation toward its maximum (1.0), and negative net input drives unit activation toward its minimum (0.0). Complete model implementation details are provided in the Appendix, along with parameter values and ranges over which the simulated effects held. Full model code is available in the supplemental materials (http://dx.doi.org/10.1037/a0027175.supp).

All of the simulations used a simple lexicon consisting of five two-letter (or two-phoneme) words. Each word was also associated with 10 semantic feature units. Four of the words were neighbors (of whatever type was relevant for the simulation); the fifth word had no neighbors. The simulations compared processing of Word 1”, a high neighborhood density word, and Word 5, the low neighborhood density word.

Simulations 1 Through 3: Word Recognition

As reviewed earlier, form neighbors tend to facilitate visual word recognition (e.g., Andrews, 1997; Yates, 2005) but have an inhibitory effect if they are higher in word frequency than the target word (Davis et al., 2009; Ferraro & Hansen, 2002; Grainger & Jacobs, 1996; Grainger et al., 1989, 1992; Grainger & Segui, 1990) or if the task is spoken word recognition (e.g., Luce & Pisoni, 1998; Magnuson et al., 2007). For these simulations, the words in the high-density neighborhood shared their first letter or phoneme and the word in the low-density neighborhood shared no letters or phonemes with any other word. To simulate processing of a printed word (Simulations 1 and 2), the units corresponding to its constituent letters were activated simultaneously, and activation was allowed to propagate through the network until one of the word units crossed a response threshold value (0.7). Simulated response time was taken to be the number of time steps needed to reach this threshold. To simulate the effect of word frequency (Simulation 2), the connection weights between the higher frequency word units and the corresponding letter units were increased. This implementation was chosen on the basis of evidence that word frequency affects the strength of the mappings represented by these connections (e.g., Dahan, Magnuson, & Tanenhaus, 2001) and because in models where such mappings are learned, the higher frequency mappings are learned more strongly (e.g., Gaskell & Marslen-Wilson, 1997; Plaut, McClelland, Seidenberg, & Patterson, 1996). To simulate processing of a spoken word (Simulation 3) the same input was presented sequentially. The first phoneme was activated for Time Steps 1 through 30, and the second phoneme was activated for Time Steps 25 through 54. The small amount of overlap was intended as a rough analog to coarticulation. The simulation results did not depend on the precise amount of coarticulatory overlap.

In Simulation 1 (Figure 3, left), the visual word with many neighbors was recognized faster than the word with fewer neighbors. As shown in the figure, visual word neighbors were weakly and transiently activated. This weak activation was too low to cause substantial lexical inhibition (see Figure 2), but the bidirectional feedback to the form layer units provided additional excitation to the shared letter unit, which facilitated target recognition. This is precisely the *gang* or *conspiracy effect* described for other IAC models (e.g., McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982; Taraban & McClelland, 1987; see also Dell & Gordon, 2003) and proposed to account for neighborhood facilitation.

The results of Simulation 2 (Figure 3, middle) were also consistent with the behavioral data: The word with high-frequency neighbors was processed more slowly than the word with equal frequency neighbors. Because of their stronger connections, higher frequency neighbors became active more quickly, reached a higher activation level, and remained active longer than equal-frequency neighbors. Importantly, reaching a higher level of activation allowed them to inhibit the target word more strongly. Because inhibition strength was a nonlinear function of activation, this inhibitory effect outweighed their increased recurrent facilitative effect on the shared letter unit. As the other simulations show next, this balance between the recurrent facilitative effects of neighbors and their lateral inhibitory effects—which hinges on their degree of activation—determines whether neighbors exert a net facilitative or inhibitory effect on processing.

The change to serial input (Simulation 3) from parallel input (Simulation 1) reversed the net effect of lexical neighbors. When the input was presented serially, lexical neighbors exerted an inhibitory effect on word processing (Figure 3, right). This comparison replicates the previously reported results from a very similar model that also showed opposite effects of neighbors for serial versus parallel input (Magnuson & Mirman, 2007). Note that

[2] Simulations 1 through 4 were concerned with interactions between the form layer and the word layer during tasks in which the semantic properties of target words were controlled and balanced, so for simplicity, the semantic layer was omitted from these simulations.

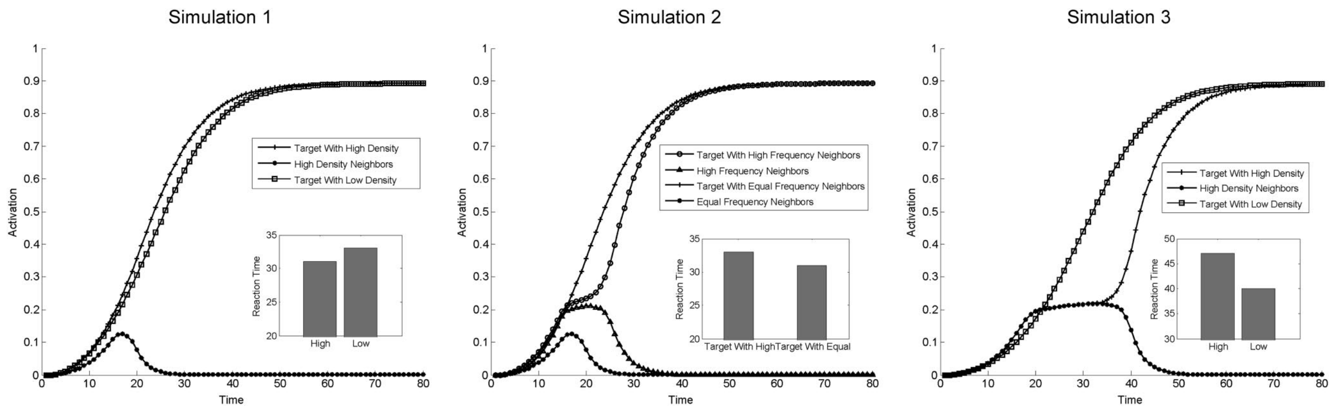

Figure 3. Results of Simulations 1 through 3. The left panel shows Simulation 1 (visual word recognition), the middle panel shows Simulation 2 (neighbor frequency effect on visual word recognition), and the right panel shows Simulation 3 (spoken word recognition). The curves represent the time course of activation for high neighborhood density target words (crosses), low neighborhood density target words (squares), and target words with higher frequency neighbors (open circles). Activation of the equal frequency neighbors (filled circles) and higher frequency neighbors (filled triangles) is also shown. The inset bar graphs show simulated word recognition reaction times based on number of processing cycles required to reach the response threshold.

neighbors were more strongly activated by serial input than parallel input (cf. filled circles in left and right panels of Figure 3). As in Simulation 2, when neighbors became more active, their inhibitory effect on the target word began to outweigh their facilitative recurrence with the form layer. In Simulations 2 and 3, the increased activation of neighbors was due to different causes—in Simulation 2 it was due to their higher frequency (stronger connections), and in Simulation 3 it was due to lack of disambiguating information during the first phoneme (the target and the neighbors had the same first phoneme). Nevertheless, the outcome was the same: The facilitative effect of neighbors became inhibitory when the neighbors were activated more strongly.

Simulation 4: Word Production

As reviewed earlier, lexical neighbors facilitate word production, making it faster and less error prone (Goldrick et al., 2010; Gordon, 2002; Middleton & Schwartz, 2010; Mirman et al., 2010; Vitevitch, 1997, 2002; Vitevitch & Sommers, 2003). To simulate this task, the model from Simulation 3 was (conceptually) run in reverse: Semantic input was presented to the word layer and allowed to propagate through the model; phoneme unit activations were treated as the model output (the same units could be considered grapheme output for simulating written word production; cf. Goldrick et al., 2010). Reaction time in word production studies is usually measured from stimulus onset to the onset of vocalization. To create a model analog of this measure, we considered the number of processing cycles from the onset of the semantic input until at least one phoneme (assumed to be the first phoneme in the vocalization) passed an activation threshold (0.7).[3]

The results (Figure 4) were consistent with the behavioral data: Word production was faster for a word with many neighbors than for a word with no neighbors. Feedback from the form layer partially activated the lexical units corresponding to the neighbors, but this activation was not enough to cause substantial inhibition of the target word. It was, however, enough to provide additional support to the shared phoneme in the target word, thus facilitating its production. One consequence of this is that the model predicts that neighbor effects should be largest on the neighbor-supported phonemes. This prediction is consistent with behavioral data showing that, even when the number of neighbors is matched, word production is faster when neighbors support more of the phoneme positions in the target word (Yates, Friend, & Ploetz, 2008). Although the effect of neighbors was largest on the shared phoneme, because of the recurrent feedback loop between phoneme and word units, the neighbor-supported phoneme unit also provided additional excitatory activation to the target word, which also facilitated activation of the other, not shared, phoneme. Simply put, facilitative neighbor effects are predicted to be strongest on the shared phonemes but are not limited to those phonemes.

The contrast between Simulations 3 and 4 (word recognition vs. word production) is consistent with simulations of a similar IAC model (Dell & Gordon, 2003). Dell and Gordon (2003) pointed out that phonological neighbors should be more strongly activated in word recognition, which is a phonologically driven task, than in word production, which is a semantically driven task. Our simulations are consistent with this claim and demonstrate that this degree of activation difference determines whether the net effect of neighbors is facilitative (weakly active neighbors) or inhibitory (strongly active neighbors).

[3] This should not be taken to mean that the model's articulatory planning is strictly phoneme by phoneme. Both output phonemes were activated in parallel, and the activation of the second phoneme lagged only slightly behind the first phoneme, reaching an average activation of 0.67 when the first phoneme reached the 0.7 threshold. This partial activation is broadly consistent with the observation that in natural speech, there is substantial coarticulation between phonemes.

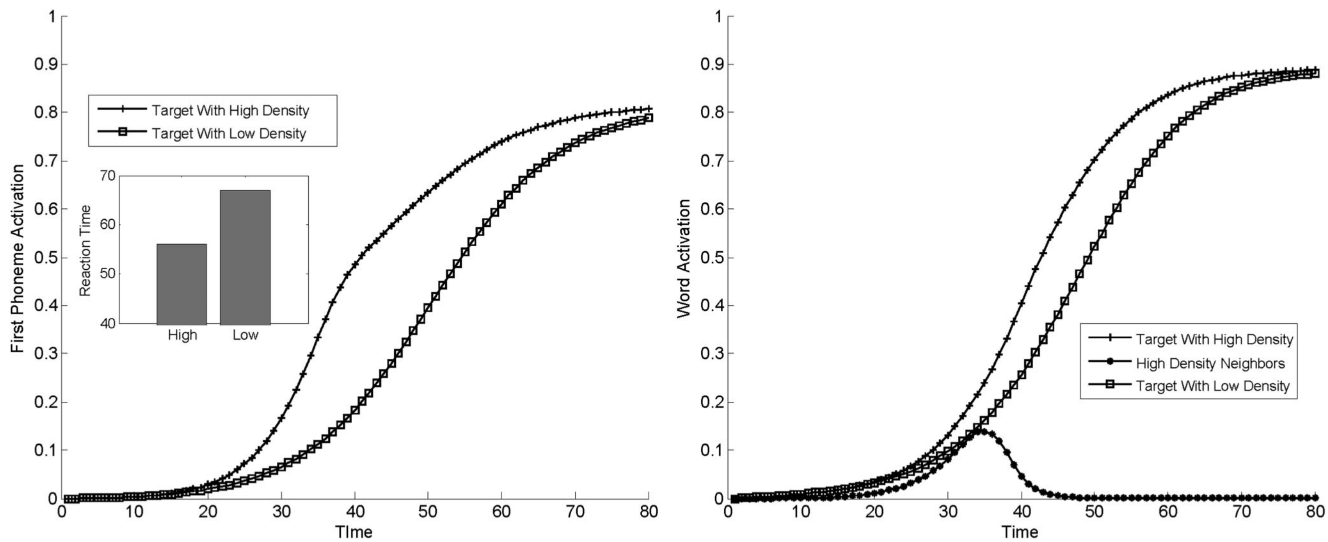

Figure 4. Results of Simulation 4: word production. The left panel shows the time course of activation for the first phoneme in a high neighborhood density word (crosses) and a low neighborhood density word (squares). The inset bar graph shows simulated word production reaction times based on number of processing cycles required for the phoneme to reach the response threshold. The right panel shows the time course of activation for target words in high-density neighborhoods (crosses) and low-density neighborhoods (squares). Filled circles indicate activation of the neighbors during processing of the high neighborhood density target word.

Simulations 5 and 6: Effects of Near and Distant Semantic Neighbors

Simulations 1 through 4 focused on effects of form neighbors and on reversals in those effects due to task or stimulus differences. In the domain of semantic neighbors, such reversals have been demonstrated within task. Specifically, both visual word recognition (Mirman & Magnuson, 2008) and word production (Mirman, 2011) were slower for words with many near semantic neighbors (concepts that share many semantic features) and faster for words with many distant semantic neighbors (concepts that share a few features). Because the materials in those behavioral experiments were matched on orthographic and phonological neighborhood, for simplicity of simulations, only the word and semantic layers were included in these simulations. To simulate word recognition (Simulation 5), input was presented directly to the word layer and allowed to propagate bidirectionally between the word layer and the semantic layer. To simulate word production (Simulation 6), input was presented to each of the semantic feature units corresponding to the target word.

Each word unit was connected to 10 semantic feature units. Lateral connections among semantic feature units were set on the basis of whether those features tended to co-occur across concepts. Feature units that never occurred together in a concept were connected by negative weights (-0.03), feature units that sometimes occurred together and sometimes separately were connected by small positive weights (0.002), and feature units that always occurred together were connected by positive weights whose magnitude was a function of the number of co-occurrences (C; e.g., Cree & McRae, 2003; Rogers & McClelland, 2004):

$$W(C) = 0.016 + 0.004 \times C.$$

Near semantic neighbors were defined as sharing eight out of 10 semantic features; distant semantic neighbors were defined as sharing four out of 10 semantic features. In the behavioral studies (Mirman, 2011; Mirman & Magnuson, 2008), the term *many* meant a quite different number for near and distant neighbors: For near neighbors, it was around four, but for distant neighbors it was over 200. To capture this distinction, the word with many near neighbors had one near neighbor, and the word with many distant neighbors had 10 distant neighbors. As in the previous simulations, the word with few neighbors had no neighbors (near or distant). For Simulation 5, because the behavioral study used a semantic task (concreteness judgment), we considered the semantic layer to be the output layer. However, because of the distributed semantic representation, there was no single unit whose activation would correspond to target (or neighbor) activation. As a result, word recognition was assessed using normalized cross-entropy error (Mirman & Magnuson, 2008; see the Appendix for more details) to measure the distance in semantic state space between the model's state and ideal states corresponding to target and neighbor words. Reaction time was measured as the number of time steps from stimulus onset (start of input to word layer) until the normalized cross-entropy error dropped below a threshold (0.2), that is, the number of time steps required for the activation pattern in the semantic layer to get close enough to the target word to consider the word to have been recognized. For Simulation 6 (word production), the model's task performance was based directly on activations of the word layer units and response time computed as the number of time steps required for the target word unit to reach the response threshold (0.7, as in the other simulations).

The simulation results (Figure 5) showed that the model exhibited opposite effects of near and distant semantic neighbors on word recognition (Simulation 5, left panels in Figure 5) and word production (Simulation 6, right panels in Figure 5). In Simulation 5, the model was slower to settle to the target representation when the target word had many near semantic neighbors than when it had none (Figure 5, top left panel) and was faster when the target word had many distant semantic neighbors than when it had none (Figure 5, bottom left panel). In Simulation 6, word activation was slower for words with many near semantic neighbors (Figure 5, top right panel) and was faster for words with many distant semantic neighbors (Figure 5, bottom right panel). As in the other simulations, strongly activated neighbors (near neighbors) had a net inhibitory effect on word recognition, and weakly activated neighbors (distant neighbors) had a net facilitative effect on word recognition.

General Discussion

Summary of Key Findings

One of the most widely agreed on principles in cognition is that multiple similar representations are activated in parallel and com-

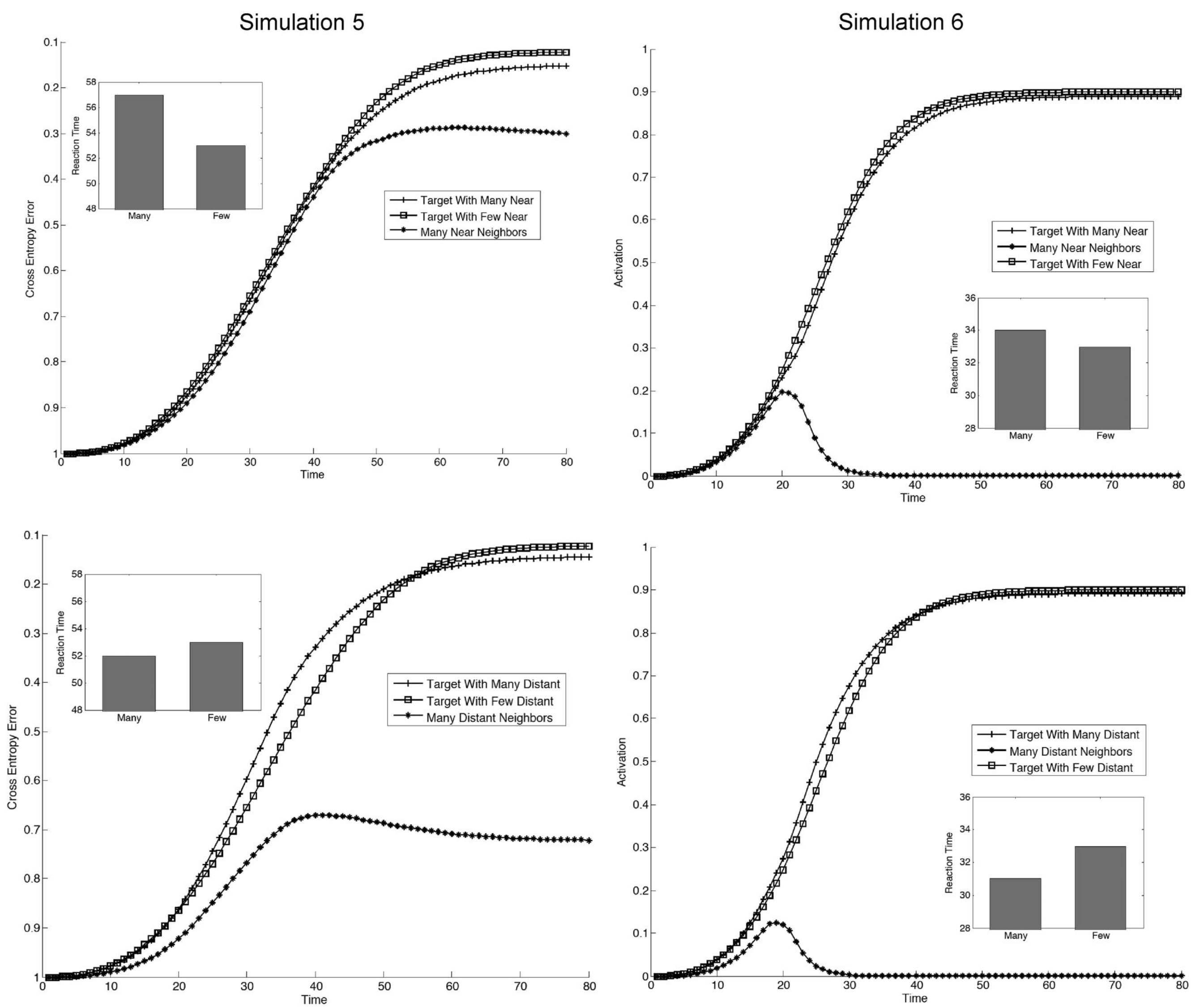

Figure 5. Results of Simulations 5 and 6: effects of semantic neighbors on word recognition (Simulation 5, left panels) and word production (Simulation 6, right panels). The top row shows effects of near neighbors, and the bottom row shows effects of distant neighbors. The curves show the time course of processing (settling in Simulation 5 and word activation in Simulation 6) for the low neighborhood density word (squares), the high neighborhood density word (crosses), and the neighbor words (circles). In the left panels, the vertical axis is reversed to reflect that lower normalized cross-entropy error corresponds to higher proximity to the target concept in semantic space. The inset bar graphs show simulated word recognition reaction times based on number of processing cycles required for the model to reach the response threshold.

pete for selection. This principle plays a key role in theories and models across a diverse set of domains, including perception (e.g., Palmeri, Wong, & Gauthier, 2004), action (e.g., Botvinick, Buxbaum, Bylsma, & Jax, 2009; Cisek, 2006), categorization (e.g., Kalish et al., 2004), memory (e.g., Polyn, Norman, & Kahana, 2009), and language (e.g., McClelland & Rumelhart, 1981). In the language domain, this principle has been studied extensively in the context of neighborhood effects—how recognition or production of a target word is affected by words that are similar to it. Orthographic, phonological, and semantic similarity have been considered in a wide variety of tasks (picture naming, word reading, word repetition, lexical decision, semantic categorization or judgment, etc.). Across these many studies, a striking pattern of consistent reversals has emerged: Given a particular task and neighbor type, the effects are quite consistent, but the direction of the effect—facilitation versus inhibition—differs across tasks and neighbor types. Although neighbor effects are one of the most robust findings in lexical processing tasks, there has been little effort to explain why the same neighbors would, for example, have facilitative effects on spoken word production, inhibitory effects on spoken word recognition, and facilitative effects on visual word recognition. Further, the accounts of individual facilitative or inhibitory effects have almost universally appealed to IAC, without addressing why the same framework would predict opposite effects in these different contexts.

Here, we have addressed this specific question using a simple implementation of the IAC framework. We deliberately chose a simple version of IAC that did not strive to capture all of the details of any one task so that we could use the same model across tasks. The simulation results captured the core qualitative patterns of orthographic, phonological, and semantic neighbor effects in word recognition and production tasks (summarized in Table 1). By using the same model with the same parameter values across simulations, we were able to extract a core computational principle that determined whether neighbor effects were facilitative or inhibitory: Strongly active neighbors had a net inhibitory effect, and weakly active neighbors had a net facilitative effect. This pattern emerged from comparisons across simulations and was not dependent on specific parameter settings (i.e., the qualitative simulation results held over fairly large changes in parameter values; see Appendix for details). This pattern is also consistent with the few previous attempts to explain why neighbors have opposite effects in different tasks (Dell & Gordon, 2003; Magnuson & Mirman, 2007).

The qualitative results demonstrated in the present simulations critically depend on the sigmoid inhibition function implemented at the word layer. This implementation was rooted in standard models of decision processes, most notably the Luce (1959) choice rule (see also Usher & McClelland, 2001), and in neural evidence (e.g., Cisek, 2006). Further, in Simulation 4 (effects of phonological neighbors on word production) and Simulation 5 (effects of semantic neighbors on word recognition), the model output was not read directly from the word layer, so idiosyncratic effects limited to the word layer could not explain those results. In other words, the sigmoid inhibition function played a critical role in the model dynamics, not just word level dynamics. Finally, when considered from the perspective of explanatory power, evidence that this single (well-motivated) processing principle can account for a large and complex set of qualitative data patterns, and do so over a relatively large parameter range, suggests that this principle may be an important aspect of cognitive processing. According to this view, the simulation results reported here serve as an existence proof demonstrating that a diverse set of findings can be explained by an underlying nonlinear (sigmoid) relationship between activation and inhibition of competing representations.

Limitations, Speculations, and Future Directions

Effects of lexical neighbors on word processing are among the most studied phenomena in lexical processing, and highly detailed models have been developed to account for individual kinds of neighborhood effects, such as orthographic neighbors in visual word recognition. However, so far, there has been little effort to develop a unified account across tasks and neighbor types. Here, we have taken a step toward developing such an account. Because the goal was to examine domain-general computational properties, we used a very simple and general model that was designed to capture the principles of IAC in a way that is consistent with existing models of written and spoken language processing (Dell et al., 1997; Jacobs & Grainger, 1994; McClelland & Elman, 1986; McClelland & Rumelhart, 1981) and general enough to be applied to very different tasks. That is, our model was meant to be the beginning of a bridge across domain-specific accounts of neighborhood effects.

Domain-specific models have provided very detailed (even item-level) accounts of neighborhood effects in their domains, but such models are limited in that they do not speak to any neighbor effects in other domains. Our model addresses effects of neighbors across domains (indeed, the model was general enough that it is possible to relabel it to be a model of object recognition, memory, categorization, or any other domain where IAC are proposed as core mechanisms), but it is limited in that it does not capture important domain-specific factors. We believe our simulations provide important complementary evidence that will help to guide the integration of domain-specific models into a domain-general account that captures both the strengths of our simple domain-general model and existing domain-specific models.

A second, related limitation is that the present simulations used very small lexicons with very simple neighbor relations. As a result, the domain generality that we observed in our simulations may have been bolstered by the similarity of the neighborhood structures that we used. Studies with larger scale models that build on our model and implement more domain-specific details are needed to investigate the balance between domain-specific versus domain-general properties of neighbor effects.

Although our model used a simple neighbor definition, the general principle that it demonstrates is applicable to effects of neighbors defined in other ways. One such case is the inhibitory effect of transposed-letter neighbors in visual word recognition (Andrews, 1996): Words like *salt*, which has the transposed-letter neighbor *slat*, are recognized more slowly than matched words that do not have a transposed-letter neighbor (such as *halt*). This pattern could arise if transposed-letter neighbors are more confusable than substituted-letter neighbors, which have facilitative effects on visual word recognition, that is, if transposed-letter neighbors are stronger competitors than substituted-letter neighbors. This would be analogous to the results of Simulations 5 and 6, which showed inhibitory effects of near semantic neighbors and facilitative effects of distant semantic neighbors, and to Simulation 2, which showed that high-frequency neighbors inhibit target word recognition because they are more active.

Another important consideration is the clustering or spread of neighbors, which has been shown to affect spoken (Chan & Vitevitch, 2009; Vitevitch, 2007) and visual (Mathey & Zagar, 2000) word recognition and word production (Yates et al., 2008). We have focused on the effects of neighbors on target word processing, but the neighbors also affect one another, which should have indirect but measurable effects on target processing. Our simulations predict that if particular patterns of neighbor clustering lead the neighbors to enhance one another's activation, then they will tend to have more inhibitory (or less facilitative) effects; in contrast, if they do not accentuate one another's activation, then their cumulative effect on the target will be more facilitative. This could explain why clustered neighbors (i.e., neighbors that are also neighbors of one another) are particularly inhibitory in spoken word recognition (Chan & Vitevitch, 2009).

A related question is whether phonotactic probability facilitates spoken word recognition, in contrast with the inhibitory effects of lexical neighbors in the same conditions (Luce & Large, 2001; Vitevitch & Luce, 1998, 1999; Vitevitch et al., 1999). Our simulations suggest that this pattern could arise if phonotactic probability is indirectly measuring distant phonological neighborhoods, that is, if high phonotactic probability words cause diffuse, weak activation of lexical representations, which would have a net facilitative effect. This would contrast with traditional lexical neighborhood metrics (e.g., the one-phoneme rule), which capture near phonological neighbors that are strongly activated and would have a net inhibitory effect. Perhaps the difficulty in replicating the phonotactic probability effect (e.g., Lipinski & Gupta, 2005; Roodenrys & Hinton, 2002; T. Strauss, personal communication, May 5, 2009) stems from it indirectly measuring distant phonological neighbors. A more direct manipulation may provide clearer insights into this issue.

The present simulations relied on a simple hard-wired model that had no learning process and no intrinsic noise during processing. Thus, the model did not address how neighborhood effects interact with learning, development, neurological impairment, and cognitive decline due to normal aging. Similarly, the model did not address errors in word recognition and production. These are interesting and important issues for future research.

Finally, the present simulations relied on qualitative, rather than quantitative, comparisons of model and behavioral data. We chose this approach because there were such striking and unexplained qualitative reversals in the literature. By choosing a simple model, we were able to use the same model to qualitatively account for the full set of behavioral data and to extract global (relatively parameter-independent) insights into the dynamics of IAC. It would be a mistake to expect such a simple model, chosen for domain generality rather than item-level specificity, to provide quantitative fits to behavioral data (see Pitt et al., 2006, for a discussion of different model evaluation methods). Rather, the basic principles explored here can form the basis for elaborated models that can and should be evaluated quantitatively.

Concluding Remarks

We have presented a series of simulations exploring whether the dynamics of IAC can account for facilitative and inhibitory effects of lexical neighbors. We chose to explore this issue in the domain of single word processing, but the computational principle is one of the most general in theories of cognition: parallel activation of multiple similar representations and competition (selection) among them. Single word processing provided an ideal context for investigating this issue because there is a large literature showing a complex pattern of facilitative and inhibitory effects, which have been robustly replicated and explained individually but not together. We used a simple IAC model that could be applied to the full range of behavioral data without changing the model architecture or parameters. The simulations showed that the complex pattern of contrasting neighbor effects boils down to a simple computational principle: Coactivated representations have both facilitative and inhibitory effects; they have a net inhibitory effect if they are strongly activated and a net facilitative effect if they are weakly activated. Because the model is so general, this insight applies to any domain where IAC are thought to be involved, that is, any mental activity involving the processing of related or similar representations in a multilevel system—which covers much of perception, cognition, and action.

References

Acha, J., & Perea, M. (2008). The effect of neighborhood frequency in reading: Evidence with transposed-letter neighbors. *Cognition, 108,* 290–300. doi:10.1016/j.cognition.2008.02.006

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419–439. doi:10.1006/jmla.1997.2558

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 802–814. doi:10.1037/0278-7393.15.5.802

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. doi:10.1037/0278-7393.18.2.234

Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language, 35,* 775–800. doi:10.1006/jmla.1996.0040

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4,* 439–461. doi:10.3758/BF03214334

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316. doi:10.1037/0096-3445.133.2.283

Benkí, J. R. (2003). Quantitative evaluation of lexical status, word frequency, and neighborhood density as context effects in spoken word recognition. *Journal of the Acoustical Society of America, 113,* 1689–1705. doi:10.1121/1.1534102

Botvinick, M. M., Buxbaum, L. J., Bylsma, L. M., & Jax, S. A. (2009). Toward an integrated account of object and action selection: A computational analysis and empirical findings from reaching-to-grasp and tool-use. *Neuropsychologia, 47,* 671–683. doi:10.1016/j.neuropsychologia.2008.11.024

Boyczuk, J. P., & Baum, S. R. (1999). The influence of neighborhood density on phonetic categorization in aphasia. *Brain and Language, 67,* 46–70. doi:10.1006/brln.1998.2049

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review, 8,* 531–544. doi:10.3758/BF03196189

Chan, K. Y., & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1934–1949. doi:10.1037/a0016902

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: A computational model. *Journal of Neuroscience, 26,* 9761–9770. doi:10.1523/JNEUROSCI.5605-05.2006

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. doi:10.1037/0033-295X.97.3.332

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of *chipmunk, cherry, chisel, cheese,* and *cello* (and many other such concrete nouns). *Journal of Experimental Psychology: General, 132,* 163–201. doi:10.1037/0096-3445.132.2.163

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science, 23,* 371–414. doi:10.1207/s15516709cog2303_4

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42,* 317–367. doi:10.1006/cogp.2001.0750

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117,* 713–758.

Davis, C. J., Perea, M., & Acha, J. (2009). Re(de)fining the orthographic neighborhood: The role of addition and deletion neighbors in lexical decision and reading. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1550–1570. doi:10.1037/a0014253

Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 9–38). Boston, MA: de Gruyter Mouton. doi:10.1515/9783110895094.9

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104,* 801–838. doi:10.1037/0033-295X.104.4.801

Dufour, S., & Peereman, R. (2003a). Inhibitory priming effects in auditory word recognition: When the target's competitors conflict with the prime word. *Cognition, 88,* B33–B44. doi:10.1016/S0010-0277(03)00046-5

Dufour, S., & Peereman, R. (2003b). Lexical competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory & Cognition, 31,* 1271–1283. doi:10.3758/BF03195810

Duñabeitia, J. A., Avilés, A., & Carreiras, M. (2008). NoA's ark: Influence of the number of associates in visual word recognition. *Psychonomic Bulletin & Review, 15,* 1072–1077. doi:10.3758/PBR.15.6.1072

Ferraro, F. R., & Hansen, C. L. (2002). Orthographic neighborhood size, number of word meanings, and number of higher frequency neighbors. *Brain and Language, 82,* 200–205. doi:10.1016/S0093-934X(02)00016-0

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 696–713. doi:10.1037/0278-7393.22.3.696

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45,* 468–492. doi:10.1006/jmla.2000.2784

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12,* 613–656. doi:10.1080/016909697386646

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28,* 501–518. doi:10.1016/0749-596X(89)90009-0

Goldrick, M., Folk, J. R., & Rapp, B. (2010). Mrs. Malaprop's neighborhood: Using word errors to reveal neighborhood structure. *Journal of Memory and Language, 62,* 113–134. doi:10.1016/j.jml.2009.11.008

Gordon, J. K. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language, 82,* 113–145. doi:10.1016/S0093-934X(02)00001-9

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29,* 228–244.

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes, 23,* 1–35. doi:10.1080/01690960701578013

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565. doi:10.1037/0033-295X.103.3.518

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics, 45,* 189–195. doi:10.3758/BF03210696

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1992). Neighborhood frequency effects and letter visibility in visual word recognition. *Perception & Psychophysics, 51,* 49–56. doi:10.3758/BF03205073

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics, 47,* 191–198. doi:10.3758/BF03205983

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), *Cambridge handbook of computational psychology* (pp. 59–100). Cambridge, United Kingdom: Cambridge University Press.

Harley, T. A., & Brown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology, 89,* 151–174. doi:10.1111/j.2044-8295.1998.tb02677.x

Jacobs, A. M., & Grainger, J. (1994). Models of visual word recognition: Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1311–1332.

Johnson, N. F., & Pugh, K. R. (1994). A cohort model of visual word recognition. *Cognitive Psychology, 26,* 240–346. doi:10.1006/cogp.1994.1008

Johnson, R. L. (2009). The quiet clam is quite calm: Transposed-letter neighborhood effects on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 943–969. doi:10.1037/a0015572

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: Knowledge partitioning and function learning. *Psychological Review, 111,* 1072–1099. doi:10.1037/0033-295X.111.4.1072

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25,* 463–492. doi:10.1080/02643290701674851

Lipinski, J., & Gupta, P. (2005). Does neighborhood density influence repetition latency for nonwords? Separating the effects of density and duration. *Journal of Memory and Language, 52,* 171–192. doi:10.1016/j.jml.2004.10.004

Locker, L., Simpson, G. B., & Yates, M. (2003). Semantic neighborhood effects on the recognition of ambiguous words. *Memory & Cognition, 31,* 505–515. doi:10.3758/BF03196092

Luce, P. A. (1986). *Neighborhoods of words in the mental lexicon: Research on speech perception* (Technical Report No. 6). Bloomington, IN: Speech Research Laboratory, Department of Psychology, Indiana University.

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes, 16,* 565–581. doi:10.1080/01690960143000137

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19,* 1–36. doi:10.1097/00003446-199802000-00001

Luce, R. D. (1959). *Individual choice behavior.* Oxford, England: Wiley.

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science, 31,* 133–156. doi:10.1080/03640210709336987

Magnuson, J. S., & Mirman, D. (2007). *Neighborhood effects in word recognition: It's not where you live, it's how you get home.* Paper presented at the 48th annual meeting of the Psychonomic Society, Long Beach, CA.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General, 132,* 202–227. doi:10.1037/0096-3445.132.2.202

Mathey, S., & Zagar, D. (2000). The neighborhood distribution effect in visual word recognition: Words with single and twin neighbors. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 184–205. doi:10.1037/0096-1523.26.1.184

McClelland, J. L. (1993). Toward a theory of information processing in graded, random, and interactive networks. In D. E. Meyer & S. Kornblum (Eds.), *Attention & Performance XIV: Synergies in experimental psychology, artificial intelligence and cognitive neuroscience* (pp. 655–688). Cambridge, MA: MIT Press.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86. doi:10.1016/0010-0285(86)90015-0

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375–407. doi:10.1037/0033-295X.88.5.375

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38,* 283–312. doi:10.1006/jmla.1997.2543

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207–238. doi:10.1037/0033-295X.85.3.207

Middleton, E. L., & Schwartz, M. F. (2010). Density pervades: An analysis of phonological neighbourhood density effects in aphasic speakers with different types of naming impairment. *Cognitive Neuropsychology, 27,* 401–427. doi:10.1080/02643294.2011.570325

Mirman, D. (2011). Effects of near and distant semantic neighbors on word production. *Cognitive, Affective, and Behavioral Neuroscience, 11,* 32–43. doi:10.3758/s13415-010-0009-7

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59,* 475–494. doi:10.1016/j.jml.2007.11.006

Mirman, D., Kittredge, A. K., & Dell, G. S. (2010). Effects of near and distant phonological neighbors on picture naming. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 1447–1452). Austin, TX: Cognitive Science Society.

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 65–79. doi:10.1037/0278-7393.34.1.65

Mirman, D., Yee, E., Blumstein, S. E., & Magnuson, J. S. (2011). Theories of spoken word recognition deficits in aphasia: Evidence from eye-tracking and computational modeling. *Brain and Language, 117,* 53–68. doi:10.1016/j.bandl.2011.01.004

Mulatti, C., Reynolds, M. G., & Besner, D. (2006). Neighborhood effects in reading aloud: New findings and new challenges for computational models. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 799–810. doi:10.1037/0096-1523.32.4.799

Newman, R. S., Sawusch, J. R., & Luce, P. A. (1997). Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 873–889. doi:10.1037/0096-1523.23.3.873

Palmeri, T. J., Wong, A. C.-N., & Gauthier, I. (2004). Computational approaches to the development of perceptual expertise. *Trends in Cognitive Sciences, 8,* 378–386. doi:10.1016/j.tics.2004.06.001

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114,* 273–315.

Pitt, M. A., Kim, W., Navarro, D. J., & Myung, J. I. (2006). Global model analysis by parameter space partitioning. *Psychological Review, 113,* 57–83. doi:10.1037/0033-295X.113.1.57

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. doi:10.1037/0033-295X.103.1.56

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review, 116,* 129–156. doi:10.1037/a0014420

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach.* Cambridge, MA: MIT Press.

Roodenrys, S., & Hinton, M. (2002). Sublexical or lexical effects on serial recall of nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 29–33. doi:10.1037/0278-7393.28.1.29

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94. doi:10.1037/0033-295X.89.1.60

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

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. doi:10.1037/0096-1523.21.4.876

Siakaluk, P. D., Buchanan, L., & Westbury, C. (2003). The effect of semantic distance in yes/no and go/no-go semantic categorization tasks. *Memory & Cognition, 31,* 100–113. doi:10.3758/BF03196086

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26,* 608–631. doi:10.1016/0749-596X(87)90105-7

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592. doi:10.1037/0033-295X.108.3.550

Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Language and Speech, 40,* 211–228.

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 735–747. doi:10.1037/0278-7393.28.4.735

Vitevitch, M. S. (2007). The spread of the phonological neighborhood influences spoken word recognition. *Memory & Cognition, 35,* 166–175. doi:10.3758/BF03195952

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9,* 325–329. doi:10.1111/1467-9280.00064

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40,* 374–408. doi:10.1006/jmla.1998.2618

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language,* 68, 306–311. doi:10.1006/brln.1999.2116

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition, 31,* 491–504. doi:10.3758/BF03196091

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15,* 971–979. doi:10.3758/PBR.15.5.971

Yates, M. (2005). Phonological neighbors speed visual word processing: Evidence from multiple tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1385–1397. doi:10.1037/0278-7393.31.6.1385

Yates, M., Friend, J., & Ploetz, D. M. (2008). Phonological neighbors influence word naming through the least supported phoneme. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1599–1608. doi:10.1037/a0011633

Yates, M., Locker, L., & Simpson, G. B. (2003). Semantic and phonological influences on the processing of words and pseudohomophones. *Memory & Cognition, 31,* 856–866. doi:10.3758/BF03196440

Yates, M., Locker, L., & Simpson, G. B. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin & Review, 11,* 452–457. doi:10.3758/BF03196594

Appendix

Model Details

The model was a simple *neural network* implementation of the IAC framework (e.g., McClelland & Rumelhart, 1981). The network was composed of simple processing units organized into layers. Each unit had an activation value and was connected to other units by weighted connections, which could be excitatory or inhibitory. At the start of each simulation, all units were initialized to the rest activation values (0) and external input was provided to the appropriate units. This activation was then propagated by the weighted connections over a series of time steps. On each time step, each unit computed its net input from all units connected to it on the basis of their activations. Specifically, the net input to unit i was

$$net_i = extinput_i + \sum_j w_{ij} a_j,$$

where w_{ij} is the connection weight from unit j to unit i, a_j is the activation of unit j, and $extinput_i$ is any external input to unit i. Every unit's activation was then updated as follows:

$$\text{If } (net_i > 0), \quad \Delta a_i = (\max - a_i) net_i - decay(a_i - rest).$$

$$\text{Otherwise,} \quad \Delta a_i = (a_i - \min) net_i - decay(a_i - rest),$$

where *max* is the maximum activation, *min* is the minimum activation, *rest* is the resting activation level, and *decay* is a constant that brings the activation of the unit back to resting activation level. In all simulations reported in this article, we choose $max = 1$, $min = 0$, and $rest = 0$. These are typical values used in many IAC network models. See Table A1 for the full list of other model parameters and their values.

At the word layer, the inhibitory weights were scaled by a sigmoid function of their activation:

$$y = \frac{15}{1.5 + e^{-\beta(x - x_0)}},$$

where increasing the value of the parameter β increases the steepness of the curve, and x_0 determines the *crossover point* (i.e., the activation value at which the inhibition weight scaling factor is halfway between its minimum and maximum). For all simulations reported here, $\beta = 35$ and $x_0 = 0.3$; the sigmoid curve corresponding to these parameter values is shown in Figure 2. This sigmoid scaling allows multiple word units to be activated initially (inhibition is weak when unit activation is low) and forces the model to rapidly settle to a single active word unit (rising activation causes a fast increase in inhibition strength; see also, Cisek, 2006).

The qualitative simulation results were quite robust over a large range of parameter values. The results required a sigmoid scaling of the inhibition strength between words, but the particular parameters of the sigmoid function were not critical. For the parameters in Table A1, almost an entire ±50% range of the given values produced all of the same qualitative patterns. One relatively sensitive parameter was the frequency scaling factor in Simulation 2

(*Appendix follows*)

Table A1
Parameter Values Used in All Simulations

Parameter	Value
Phoneme/letter to word excitation	0.1
Word to phoneme/letter excitation	0.1
Word to semantics excitation	0.03
Semantics to word excitation	0.03
Word to word inhibition	0.04
Phoneme/letter decay	0.01
Word layer decay	0.01
Semantic layer decay	0.05

Note. The excitation and inhibition parameters refer to connection weights.

(increased by 36%), which determined the strength of letter-to-word excitation weights for higher frequency words relative to lower frequency words. This parameter needed to be balanced with the model's other excitatory and inhibitory weights to achieve the correct pattern of recurrent facilitative effects and lateral inhibitory effects. Note that this was a matter of balance, not absolute value: For any choice of the parameters in Table A1, there was a range for this scaling factor that produced the reported result, but this range was specific to those other parameter values.

Cross-Entropy Calculation

In Simulation 5, the output was read from the distributed semantic representation, which could not be evaluated on the basis of activation of a single unit. We used cross-entropy error (CEE) to assess the distance between the activation pattern in the semantic layer and the target activation pattern. The CEE function is

$$CEE = -\sum_j d_j \ln(s_j) + (1-d_j)\ln(1-s_j),$$

where d_j is the target activation for the jth semantics unit, and s_j is the current activation for that unit. This raw CEE value was normalized (divided by the maximum CEE for each item) to remove the effects of model starting state and map the values into the same 0 to 1 scale as the other simulations. Conceptually, normalized CEE represents the proportion of the distance between the model's starting state and the target state that the model has traversed.

Received September 6, 2011
Revision received December 28, 2011
Accepted December 28, 2011 ■

Psychological Review

0033-295X/12/$12.00 DOI: 10.1037/a0030049

Correction to Chen and Mirman (2012)

In the article "Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors" by Qi Chen and Daniel Mirman (*Psychological Review*, 2012, Vol. 119, No. 2, pp. 417-430), incorrect decay rates were provided in the Appendix Table A1. The corrected parameter values are listed below:

Table A1
Parameter Values Used in All Simulations

Parameter	Value
Phoneme/letter to word excitation	0.1
Word to phoneme/letter excitation	0.1
Word to semantics excitation	0.03
Semantics to word excitation	0.03
Word to word inhibition	0.04
Phoneme/letter decay	0.02
Word layer decay	0.02
Semantic layer decay	0.1

Note. The excitation and inhibition parameters refer to connection weights.

There was also an error in the model code provided in the supplemental materials. Updated model code is available at: http://sites.google.com/site/neighbormodel/model.

DOI: 10.1037/a0030049