Hemispheric Specialization and Independence for Word Recognition:

A Comparison of Three Computational Models

Scott A. Weems and James A. Reggia

University of Maryland at College Park

January 29, 2004

Acknowledgements: This investigation was supported by NIH awards T32 DC00061 and RO1 NS35470.

Corresponding author:
Scott Weems, Ph.D.
Department of Computer Science
University of Maryland
A.V. Williams Building
College Park, Maryland 20742
email: sweems@cs.umd.edu
phone: (301) 405-2670
fax: (301) 405-6707

ABSTRACT

Two findings serve as the hallmark for hemispheric specialization during lateralized lexical decision. First is an overall word advantage, with words being recognized more quickly and accurately than non-words (the effect being stronger in response latency). Second, a right visual field advantage is observed for words, with little or no hemispheric differences in the ability to identify non-words. Several theories have been proposed to account for this difference in word and non-word recognition, some by suggesting dual routes of lexical access and others by incorporating separate, and potentially independent, word and non-word detection mechanisms. We compare three previously proposed cognitive theories of hemispheric interactions (callosal relay, direct access, and cooperative hemispheres) through neural network modeling, with each network incorporating different means of interhemispheric communication. When parameters were varied to simulate left hemisphere specialization for lexical decision, only the cooperative hemispheres model showed both a consistent left hemisphere advantage for word recognition but not non-word recognition, as well as an overall word advantage. These results support the theory that neural representations of words are more strongly established in the left hemisphere through prior learning, despite open communication between the hemispheres during both learning and recall.

Keywords: neural networks, hemispheric independence, hemispheric specialization, word recognition, lexical decision, lateralization.

INTRODUCTION

Despite prominent connectivity between the left and right hemispheres, asymmetries in their cognitive functionality are well established. For over a century, researchers have observed both left and right hemisphere specialization for a variety of tasks, often through examination of remaining function following lateralized brain damage (for review, see Hugdahl & Davidson, 2003). For example, recent research shows a loss of semantic comprehension (Hagoort, 1998) and syntactic ability (Grodzinski, 2000) following left hemisphere insult, indicating an important role for the left hemisphere in these tasks. Yet we know that both hemispheres display a wide range of cognitive abilities, with language deficits often being manifest even with loss in the non-dominant hemisphere. For example, both left and right hemisphere loss have been shown to disrupt readers' ability to comprehend complex syntax (Caplan, Hildebrandt, & Makris, 1996), as well as their ability to perform lexical ambiguity resolution (Grindrod & Baum, 2003).

While the above examples involve special populations with localized damage to one or both cerebral hemispheres, controlled psychological tasks may also uncover hemispheric asymmetries. The lexical decision task, which calls on an observer to identify the lexical category (word or non-word) of a briefly presented stimulus, often shows a right visual field (RVF) advantage (Bradshaw & Gates, 1978; Chiarello, 1985; Leiber, 1976). Emotion perception and metric spatial processing, both believed to be right hemisphere specialized, are also best when lateralized to the left visual field (LVF) compared to the RVF (Banich & Federmeier, 1999; Christman & Hackworth, 1993). These visual field advantages indicate hemispheric specialization of the contralateral hemisphere due to organization of the visual system, with a RVF advantage indicating a left hemisphere specialization and a LVF advantage indiciating a right hemisphere advantage. These effects are particularly relevant because such stimulus lateralization is in many ways less straightforward than neuropsychological approaches. We know that interhemispheric transfer occurs at least as early as V1 (Houzel, Carvalo, & Lent, 2002), and that visual information is shared between the hemispheres even when the target stimulus is lateralized. Thus, such manipulations can, at best, provide a conservative measure of asymmetric cognitive abilities (for review see Zaidel, 1979).

The left hemisphere specialization for word recognition may be due to any number of cognitive factors. Orthographic, phonological, and semantic characteristics of a word may all affect fluency of its retrieval, as evidenced by both behavioral and physiological research (see Binder et al, 2003 for review). Several models have been proposed to explain psycholinguistic effects on word retrieval, and how they may affect the two cerebral hemispheres differently. Chiarello (1988) describes that lexical decision is characterized by pre-lexical, lexical, and post-lexical processing stages. The pre- and post-lexical stages involve stimulus encoding and response making, respectively, and are not believed to differ between the cerebral hemispheres. However, the lexical stage involves matching the lexical stimulus with stored entries, termed lexical access, an event that may well be left hemisphere specialized (Zouridakis, Simos, Breir, & Papanicolaou, 1998). It is during lexical access that semantic and phonological characteristics have their most clear impact, perhaps in different ways for the two hemispheres. For example, words that have been semantically primed may be easier to access, and it has been suggested that the right hemisphere possesses a more diffuse semantic association network (Burgess and Simpson 1988; see Coney 2002 for a review). It is conceivable that such semantic asymmetries may impact word recognition, even for unprimed words, although it is difficult to propose how this might occur for lexical decision. Phonological processing has also been demonstrated to differ between the hemispheres, with only the left hemisphere demonstrating a capability for phonological output, as described next.

Many models of word recognition have posited dual routes involved in lexical access (Ellis & Young, 1988; Ellis, Young, & Anderson, 1988; Young & Ellis, 1985). The first route is often equated with lexical access, as it involves matching the observed lexical stimulus with known entries in search for a match. The second route, however, is not lexical in nature and involves orthographic to phonemic translation. This translation, which produces a phonological output for the observed stimulus, is believed to be available only to the left hemisphere, thereby providing it with a processing advantage. Numerous forms of dual-route models have been proposed (see Coltheart 1993 for review), although mixed support for the use of dual routes in lexical decision has been observed (Iacoboni & Zaidel, 1996). Seidenberg and McClelland (1989) observed that dual lexical and phonemic routes are not necessary to show many of the phonological effects common with word recognition tasks. Other research concerning different aspects of linguistic ability such as verb tense translation (Joanisse & Seidenberg, 1999) and non-word pronunciation (Seidenberg, Plaut, Petersen, McClelland, & McRae, 1994) have also shown that connectionist models can perform as well as multiple system approaches (e.g. dual routes) at predicting human language performance.

One difficulty with the interpretation of hemispheric asymmetries is that it is not entirely clear what some of the terminology used to describe them means. For example, even though most neurolinguists agree that language is specialized to the left hemisphere, there is less agreement regarding the relationship between specialization and interaction. Specialization may imply dominance, with tasks such as speech production being performed solely by the left hemisphere. With such dominance, only one hemisphere is capable of performing the given task, and the second non-dominant hemisphere must take a secondary role during processing. Others take specialization to mean greater facility of one hemisphere or the other, with no limitations on which hemisphere takes control during task performance (Zaidel, Clarke, & Suyenobu, 1990). One difference between these views is apparent: the first assumes that for a lateralized cognitive task stimulus information must undergo interhemispheric transfer when presented to the non-dominant hemisphere, whereas the second makes no such assumption. Thus, although both theories acknowledge a certain degree of hemispheric specialization, they differ on the issue of hemispheric independence. The first relates greater left hemisphere ability with dependence of the right hemisphere on left hemisphere resources for successful completion (low independence), whereas the second acknowledges greater left hemisphere ability while retaining the option of right hemisphere control (independence).

There have been several models proposed for how hemispheric specialization and interaction relate, particularly in the field of language (Zaidel, Clarke, & Suyenobu, 1990; see Table 1 for a review). First, the *callosal relay model* assumes that the left hemisphere is specialized for language, and also that the right hemisphere is dependent on the left for successful linguistic processing. The term “callosal relay” is used because it assumes stimulus information is “relayed” from the right to the left hemisphere when information is presented to the LVF, thereby allowing the dominant left side to receive all input and control processing. Evidence for this model comes not only from numerous historical studies showing language loss following left hemisphere damage, but also from fMRI studies showing strong left hemisphere lateralization during lexical decision (Calandra-Buonaura et al., 2002). Calandra-Buonara et al (2002) observed that left hemisphere activation, and not right hemisphere activation, was strongly associated with successful lexical decision performance. Similar findings come from fMRI studies of phonological fluency (Alvarez-Linera et al., 2002), and phonetic and semantic analysis (Binder et al., 1997), supporting a left hemisphere dominant view.

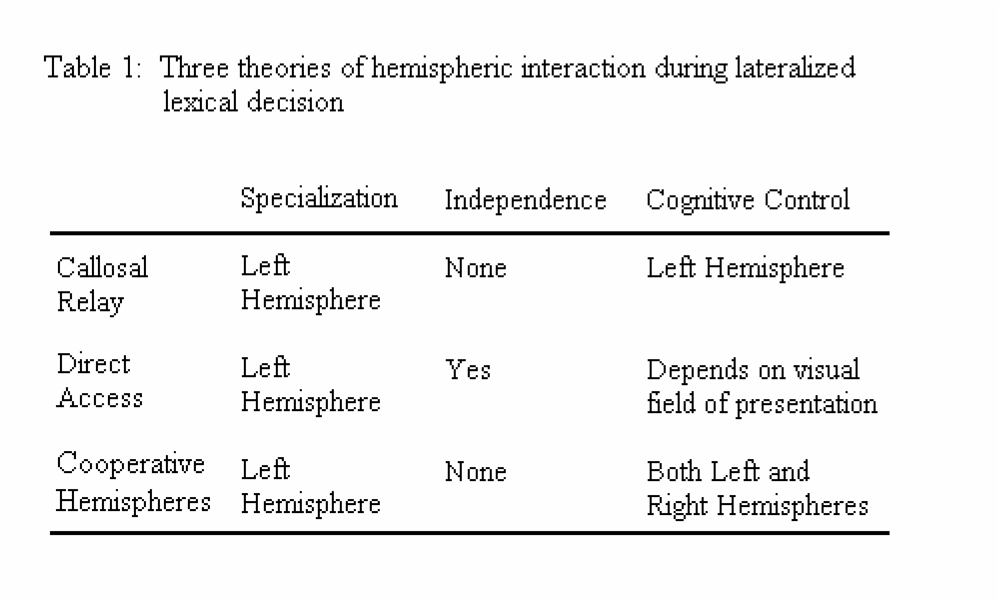

Table 1: Three theories of hemispheric interaction during lateralized lexical decision

	Specialization	Independence	Cognitive Control
Callosal Relay	Left Hemisphere	None	Left Hemisphere
Direct Access	Left Hemisphere	Yes	Depends on visual field of presentation
Cooperative Hemispheres	Left Hemisphere	None	Both Left and Right Hemispheres

Lateralized lexical decision has also uncovered evidence for the callosal relay of linguistic information between the hemispheres. Lateralized lexical decision involves a subject identifying the lexical category of a stimulus (word or non-word) presented either to the left or right of fixation, thus ensuring that it is contained within a single visual field and sent exclusively to a single hemisphere for processing (right hemisphere for LVF and left hemisphere for RVF). Mohr, Pulvermuller, and Zaidel (1994) showed that lateralized lexical decision performance was best when a copy of the target stimulus was presented to each visual field, an effect that could only be expected if the hemispheres are not independent processors. Iacoboni and Zaidel (1996) also showed that the inclusion of distracter stimuli in the visual field opposite the target reduced word recognition ability. Although distracter stimuli impaired both LVF and RVF word recognition ability, the effect was much stronger for LVF targets, as predicted by the callosal relay model. The stronger effect of RVF distracters on LVF targets supports the callosal relay model because this model requires left hemisphere involvement when the stimulus is presented to the right hemisphere (LVF): RVF distracters should, and do, have the most disruptive effect. These results do not explicitly support a callosal relay model as much as a model of hemispheric dependence (in whatever form); however they provide an important indication that resources are shared between the hemispheres during word recognition.

In contrast, the *direct access model* assumes both hemispheric specialization and independence. Thus, poor lateralized LVF performance for language tasks is not due to loss from having to transfer stimulus information from the right to the left hemisphere, but instead to reduced efficacy of the right hemisphere in performing the processing itself. For the direct access model, it is assumed that the hemisphere directly receiving the stimulus information is the one making the response, with little involvement from the contralateral hemisphere. Although it would be difficult to assert that the hemispheres work in complete isolation during word recognition, there is still good reason to believe that they retain largely independent processing abilities. The strongest proof of this comes from split-brain research showing that each hemisphere is separately capable of perceiving, reasoning, and acting without participation from the other (Bogen, 2000; Zaidel et al., 1990). It may be that the open connectivity between the hemispheres in the normal brain makes it difficult to identify these independent processing capabilities, which become apparent in the split-brain.

Although the direct access and callosal relay models have been successful in exploring the role of hemispheric independence, they both imply one hemisphere or the other “takes charge” of any given cognitive task. It is at least equally plausible that the two hemispheres work jointly towards the processing of any stimulus, such that left hemisphere specialization does not necessarily indicate a greater ability to independently perform a given cognitive task but rather a processing advantage when greater left hemisphere involvement is fostered. Visual field asymmetries may therefore tap asymmetries in hemispheric representation or participation, rather than differences in their capability to individually process a stimulus to completion. We term this model the *cooperative hemispheres model* because hemispheric independence (direct access) or dependence (callosal relay) is traded for a cooperative, dynamic relationship based on bilateral lexical representation.

The cooperative hemispheres model is often addressed in the context of neural assemblies representing lexical entries taking residence in both hemispheres. Pulvermuller and Mohr (1996) describe how bilateral neural assemblies may demonstrate left hemisphere processing advantages, as stimuli directed to that hemisphere are likely to ignite a greater portion of the assembly and therefore activate the representation. As such, the hemispheres are neither fully independent nor specialized, and processing involves an interactive process that may (or may not) show asymmetries based on the form of neural representation itself.

If words are represented jointly between the hemispheres, then the neural activity which represents them must be “bound” such that activity in different cortical areas, or in opposite hemispheres, forms a coherent assembly. Such binding may take place through synchronous or coherent neural firing, as explored through electroencephalographic (EEG) recording. Engel, Konig, Kreiter, and Singer (1991) were the first to observe synchronization of neural activity between the hemispheres for the processing of low-level stimuli in the cat. Subsequent work confirmed coherence of neural firing between the hemispheres during more complex cognitive processing in humans, such as verbal and spatial processing (Corsi-Caberra, Gutierrez, Ramos, & Arce, 1988), visual comparison (Knyazeva et al., 1999), and object recognition (Mima, Oluwatimilehin, Hiraoka, & Hallet, 2001). For language, EEG coherence increases have been associated with the processing of certain lexical stimuli (Weiss & Mueller, 2003), supporting a view of bilateral neural assemblies for lexical entries. If such neural assemblies are widely distributed, comprising neurons in multiple cortical areas and even between the hemispheres, then measuring laterality should be accomplished not by guessing which hemisphere "owns" a cognitive process, but instead by observing the relative balance of neural contributions within the hemispheres for successful performance (Pulvermuller & Mohr, 1996). Recent techniques such as fMRI provide promising approaches to this goal, and indeed have shown a relative lack of right hemisphere involvement in lexical decision (Calandra-Buonaura et al., 2002), counter to a cooperative hemispheres view.

In the current paper we use a different approach to exploring hemispheric interaction: neural network modeling. One benefit of neural network modeling is that computational models may be defined based on theories of human cognition, and then tested to measure relative success of the cognitive (i.e., theoretical) models upon which they are based. For example, we may define computational models (i.e., neural networks) based on the direct access and callosal relay theories of hemispheric interaction and train the networks to learn under these different architectures while measuring the efficacy of each. As most previous bi-hemispheric neural network research has focused on the mechanisms by which activity is shared between the hemispheres rather than the functional relevance of such transfer (Anninos, Argyrakis, & Skouras, 1984; Anninos & Cook, 1988; Cook & Beech, 1990; Reggia & Levitan, 2003; Ringo, Doty, Demeter, & Simard, 1994), this proves to be a fruitful area of current and future research. However, special concern must be made to ensure that computational models adequately represent the cognitive theories for which they stand, because without such accurate representation results from the computational models become less informative. In the current study, we recognize that some aspects of the cognitive theories regarding hemispheric specialization and interhemispheric transfer may not be completely addressed by the neural networks tested. However, the primary features of the theories are represented (e.g. asymmetric transfer, independent hemispheric output, etc…), as discussed below. Also, the different neural networks are identical except for these key features, thus ensuring other aspects of the computational models are not the source of differences in their ability to predict actual behavioral data.

Past studies using neural networks have explored the importance of hemispheric interaction, particularly in the field of language. For example, previous studies have shown that lateralization emerges in computational models during such tasks as letter identification under a variety of circumstances in which the left hemisphere shows a processing advantage (Shevtsova & Reggia, 1999; Reggia and Levitan, 2003). Specifically, left hemisphere specialization was shown to arise through larger left hemisphere size, greater cortical excitability, or a faster learning rate. Reggia and Levitan (2003) also observed that the greatest laterality occurred when models used inhibitory callosal connections, supporting the assertion that the corpus callosum is functionally inhibitory in nature (Cook, 1986)**.** Similar findings have been observed in other hemispheric models, including of phoneme sequence generation (Reggia, Goodall, & Shkuro, 1998, Shkuro et al, 2000) and dual spatial/linguistic recognition (Howard & Reggia, In press).

The current paper builds upon such previous neural network research to examine the functional significance of hemispheric asymmetries and interactions during word recognition. We use a neural network to simulate word recognition, without making any assumptions about the cognitive mechanisms by which such a task is performed in the actual brain; the only manipulation is how the hemispheres are allowed to interact. Thus, our model represents the simplest possible approach to the problem, with a single neural network being composed of two hemispheres tasked with storing, and later recognizing, a corpus of words. Three different computational models are tested, each based on one of the cognitive theories of hemispheric interaction summarized above: the callosal relay, direct access, and cooperative hemispheres theories. Although each theory/model assumes left hemisphere specialization for word recognition (Table 1), they take different perspectives regarding hemispheric independence (i.e., how the hemispheres interact). The results predicted by each model are compared to those observed in human performance for the same task to gain a beter idea for how the two hemispheres interact during normal (human) word recognition.

METHODS

Model Description

Three computational models are compared, a callosal relay model, a direct access model, and a cooperative hemispheres model. As shown in Figure 1, each consists of an input layer, two association layers (LH and RH, representing left and right hemispheric regions, respectively) and one or more output layers. The input layer is a three by five grid, with rows representing each letter and columns representing the five different possible graphemes for each letter position. The input layer is fully connected to both association layers, so during training both left and right association layers receive the same input (as justified below).

Each association layer consists of a 10 by 10 array of elements, connected in the following manner. Each node has a positive connection to its immediate neighbors with strength w_e, and an inhibitory connection to all second order neighbors (those adjacent to immediate neighbors) with strength w_i. Additionally, each node shares a positively weighted connection with its homotopic partner node in the opposite hemisphere, as well as that contralateral node's immediate neighbors (connection weight w_x). This arrangement incorporates a "Mexican-hat" topography of shared intra-hemispheric connectivity, while ensuring that activity patterns are also shared inter-hemispherically.

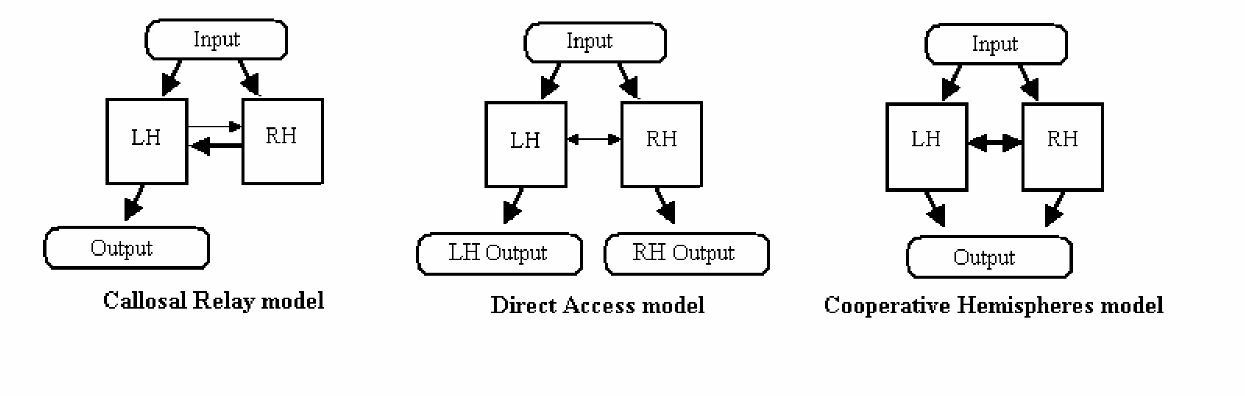

Figure 1. Three computational models are compared, the callosal relay, direct access, and cooperative hemispheres models. The callosal relay model has strong right to left, but minimal left to right connectivity, and response output only from the left hemisphere. The direct access model has minimal connectivity between the hemispheres, and separate left and right hemispheric outputs. Finally, the cooperative hemispheres model has strong interhemispheric connections, as well as a single output receiving connections from both left and right hemispheric association areas.

Activity s_{ij} for each association layer node ij is in the interval [0,1] and based on the following dynamics

$$\frac{ds_{ij}}{dt} = C_s \cdot s_{ij} + C_e \cdot \left(1 + e^{-\gamma \cdot h_{ij}}\right)^{-1}$$

where h_{ij} is the cumulative linearly-weighted activity reaching hemisphere association node ij, the parameter γ is the gain of the sigmoid function, and $C_e>0$ and $C_s<0$ are cortical excitivity and self-inhibition parameters, respectively. The activation dynamics are simulated numerically using an Euler method with a time step of *dt*=0.1. Other parameters C_p, C_{loc}, and C_x multiply the activity reaching node ij from the primary layer, neighboring intra-hemispheric, and homotopic inter-hemispheric sources, respectively, to determine the value of h_{ij}. The value for C_x, which controls relative contribution from inter-hemispheric transfer of activity, differs between the three models to control for the level of hemispheric interaction. As shown in Table 2 and as illustrated in Figure 1, C_x is minimal for left to right transfer, but significant for right to left transfer, in the callosal relay model. It is minimal in both directions for the direct access model (i.e., very little inter-hemispheric transfer of activity), but is more significant for the cooperative hemispheres model.

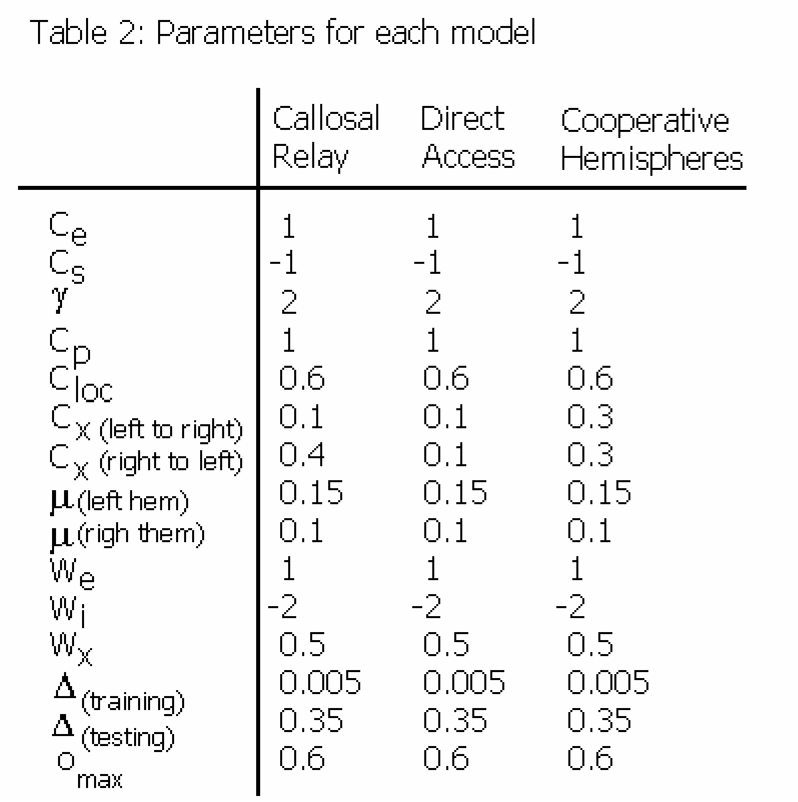

Table 2: Parameters for each model

	Callosal Relay	Direct Access	Cooperative Hemispheres
C_e	1	1	1
C_s	-1	-1	-1
γ	2	2	2
C_p	1	1	1
C_{loc}	0.6	0.6	0.6
C_x (left to right)	0.1	0.1	0.3
C_x (right to left)	0.4	0.1	0.3
μ (left hem)	0.15	0.15	0.15
μ (righ them)	0.1	0.1	0.1
W_e	1	1	1
W_i	-2	-2	-2
W_x	0.5	0.5	0.5
Δ (training)	0.005	0.005	0.005
Δ (testing)	0.35	0.35	0.35
o_{max}	0.6	0.6	0.6

The three models also differ in their connectivity to the output layer(s), a 60 node vector representing each of the 60 words in the training vocabulary (see Figure 1). The callosal relay model uses a single output layer to which the left but not right hemisphere is fully connected. Thus, only activity in the left hemisphere association area directly contributes to the output response. The direct access model includes two separate output vectors, identical except that one receives connection from the left hemisphere (left hemisphere output) and one from the right hemisphere (right hemisphere output). It is conceivable that for the direct access model the left hemisphere and right hemisphere outputs could produce conflicting outcomes, although only a single output is actually chosen based on which passes threshold first, as described below. Finally, the cooperative hemispheres model includes a single output vector that receives full connectivity from both hemispheric association areas.

Input stimuli presented to each model propagate within the recurrently connected association layers until a stable activity pattern is achieved for both hemispheres. Once activity settles, defined as $\sum\left(\left|s^{LH,t} - s^{LH,t-1}\right| + \left|s^{RH,t} - s^{RH,t-1}\right|\right) < \Delta_{training}$, then activity passes to the output layer. Each output node takes a value between 0 and 1 based on the logistic function $s_{ij} = \left(1 + e^{-\gamma \cdot h_{ij}}\right)^{-1}$ where h_{ij} is the input activity reaching output node s_{ij}, and the output node with the greatest value is designated the winner. For the direct access model, there are two output layers and so there were a total of (60 x 2) 120 possible nodes, any one of which could be a winner for any particular stimulus. Although there were two output layers, only a single node that reached threshold first was identified as winner for the direct access model, and that node could be from either the left or right hemisphere output.

All weights were initialized randomly between 0.0 and 0.1 before training. Learning in all three of the computational models was instantiated using a backpropagation learning rule, with w_{ji} being the connection weight from node i to node j, a_i being activity in node i and a_j being activity in node j:

$$\Delta w_{ji} = \mu \cdot \delta_j \cdot a_i$$

where $\delta_j = \gamma \cdot a_j \cdot (1-a_j) \cdot (t_j - a_j)$ for output nodes

and $\delta_j = \gamma \cdot a_j \cdot (1-a_j) \cdot (\sum_k \delta_k \cdot w_{jk})$ for association nodes j, summing over output nodes k

The derivative of the logistic equation embedded in the parameter δ for association nodes is appropriate since the weight changes occur only at equilibrium (i.e. after activity has settled in the cortical layers), when $s_{ij} = (-C_e/C_s) \cdot (1 + e^{-\gamma \cdot h_{ij}})$. Weights between input and association layers, as well as between association and output layers, were adjusted after activity settled. Activity was considered to have reached a steady state (i.e. settled) once activity change in a single time step dropped below threshold parameter $\Delta_{training}$. To incorporate a left hemisphere advantage, the learning rate μ for the left hemisphere was 50% higher than for the right hemisphere in all three models. There is good reason to believe that neurotransmitter concentrations differ between the cerebral hemispheres (Tucker, 1987; Tucker & Williamson, 1984). Such differences may lead to asymmetries in synaptic plasticity, and therefore hemispheric differences in learning. Although little is known about learning differences between the hemispheres, previous work has shown such learning rate asymmetries lead to hemispheric specialization similar to that which occurs in the human brain (e.g. Shevtsova & Reggia, 1999).

It would also be possible to vary other aspects of the model in order to set a left hemisphere advantage. Shevtsova and Reggia (1999) observed that greater cortical excitability, as well as larger network size, also provide processing benefits. Although it would have been possible to vary both of these aspects of the current models, we chose to focus on learning rate for three reasons. First, as described above, an advantage in learning provides a simple and intuitive way of exploring the left hemisphere advantage. It is easy to understand, and also to predict, the outcome from having a single hemisphere being a faster learner, and this advantage provided a simple way to explore left hemisphere processing advantages in word recognition Second, differences in cortical excitability and network size differences are not as easily instantiated using bi-hemispheric neural networks, as each introduces problematic asymmetries in activity levels that make performance asymmetries more difficult to identify. We hope to explore potential solutions for examining such asymmetries in the future. Third, past modeling has suggested that asymmetric hemispheric learning speed can be viewed as a common factor leadning to lateralization (i.e. it is the critical factor even if the hemispheric asymmetry used in a model is size or excitability (Reggia et al, 1998).

All parameters were the same across computational model type, except for those relating to interhemispheric transfer of activity (as described above). Cortical excitability (C_e=1) and self inhibition (C_i=-1) constants were matched with Shevtsova and Reggia's (1999) model of letter recognition, from which this model was loosely derived. Gain of the sigmoid function was set to two, and based on previous experience this value proved adequate to ensure rapid (but not too rapid) learning. The weight values for intrahemispheric excitatory (w_e=1) and inhibitory connections (w_i=-2) were chosen to ensure a strong effect of lateral inhibition while ensuring that activity in the entire network did not become suppressed by overwhelming inhibitory influences. Inter-hemispheric connection weights (w_i=.5) were set at half local excitatory connections to ensure that inter-hemispheric connections did not overpower local excitatory ones. A search of the parameter space for each of these values was conducted, and it was determined that only alteration of the interhemispheric connection weight values (w_i) significantly impacted performance of the models. If w_i was set above approximately 0.8, inter-hemispheric activity overwhelmed local intra-hemispheric influences making results uninterpretable.

A total of 1000 training epochs were conducted, with each of the 60 word stimuli presented to both hemispheres/association areas in a randomly permuted order. It was chosen to use bilateral presentation during training because we wanted to foster bilateral representation of the target stimuli. We recognize that bilateral lexical representation is a controversial topic, as is the nature of storage for lexical stimuli given the splitting of the visual fields. There is good evidence that, although the visual fields overlap some due to organization of ganglia cells in the retina (Reinhard & Trauzettel-Klosinski, 2003), both hemispheres receive only part of the input from a centrally fixated word (see Shillcock et al, 2000, for review). Thus, normal reading should involve splitting of a word between the left and right visual fields, and consequently the left and right hemispheres. However, it is an assumption of this paper, as well as those addressed earlier regarding the different cognitive models of hemispehric interaction, that lexical representation consists of a complete and intact entry in one or both hemispheres. Although this sets aside the concern regarding the splitting of visual information between the visual fields during centralized word viewing, it does allow for an examination of joint hemispheric lexical representation. It should be recognized that there is controversy regarding how words are processed during centralized, as well as lateral fixation, and that it is not assured that both hemispheres receive complete stimulus information during either presentation design. However, the goal of the current study was not to examine how words are learned during centralized or lateral presentation, but rather to learn more about how lexical entries are accessed in the two cerebral hemispheres.

A total of 10 simulations were conducted for each computational model. For each simulation weights were initially randomized, and words were presented in a different randomly permuted order for each training epoch. Appendix A includes a list of all 60 words, as well as the 60 non-words, presented during the final testing session (described below). Words were used in both training and testing, non-words were used only during testing. Words were, in effect, matrices representing possible consonant-vowel-consonant (CVC) trigrams. Row one represents all possible initial consonants, the second row prepresents all vowels, and the third row the final consonants. A total of 60 legal words, as well as 60 non-words were taken from the 125 possible CVC combinations. The bigram frequency of each word and non-word was also computed and analyzed, as discussed in the Results section.

Experimental Testing

Table 2 presents a list of all parameters used in the testing. The testing phase differed from the training phase in that stimuli were lateralized to represent unilateral visual field presentation, as is done experimentally. Thus, stimulus information was preferentially directed to one hemisphere or the other in order to identify hemispheric asymmetries. This was accomplished by presenting each stimulus in full strength to the target hemisphere (left hemisphere for RVF trials and right hemisphere for LVF trials), but only at one half strength to the contralateral hemisphere. This was necessary to ensure that one hemisphere received the bulk of stimulus information, while recognizing that at this high level of processing some interhemispheric transfer must have already taken place. It is also important to note that the current model makes no attempt to model early processes in the visual system, as the input layer is assumed to have already assembled features to the level of graphemic representation.

For each model, activity was passed to the output layer for each time step and the output node with the maximum activity level was determined. When that value passed a critical value o_{max}, then a word response was made. Also, whenever total change in activity within the association areas dropped below $\Delta_{testing}$, over a single time step a non-word response was made. These two outcomes represent the cognitive equivalents of the knowing state (word response) and cessation of the lexical search (non-word response). For the direct access model, whenever an output node in either output vector reached o_{max} a word response was made, regardless if it was the left hemisphere output layer or the right hemisphere output layer. Thus, it is possible for the left hemisphere to make a response, even when the right was initially tasked with recognizing the stimulus (LVF presentation). However, as discussed in detail in the Results section, in no case did one hemisphere respond to input presented to the contralateral hemisphere.

During testing the 60 words, as well as 60 novel non-words, were presented. Just as with behavioral lexical decision experiments, any word stimulus which led to a word response was deemed correct, even if a different word node than that presented became active past threshold. Similarly, any non-word stimulus which led to a post-threshold output node response was deemed incorrect. Response latencies, which were counted as the number of cycles before a response was made, were calculated only for correct trials. Again, this parallels behavioral experiments which traditionally show word advantage in response latency for accurately identified words.

The parameter $\Delta_{testing}$, was defined as the minimum total change in activity in the left and right hemispheres before a non-word response is made. In effect, for each input to the model two processes competed to reach completion first. Either an output node passed threshold o_{max} or the change in activity among all left and right hemisphere nodes dropped below threshold

$\Delta_{testing}$. These two parameters were carefully chosen such that response latencies for words and non-words were comparable, and it was observed that varying each had nearly identical effects on all three models: Lowering o_{max} and raising $\Delta_{testing}$ introduces a non-word bias, with the opposite changes introducing a word bias. However, it proved difficult to test any of the models with $\Delta_{testing}$ so high that mean non-word latency was greater than that for word latency, because as this occurs word accuracy falls rapidly. Consequently, although a comparison of mean word and non-word latency results are beyond the scope of this paper, it is worthy of note that it is an inherent design constraint that non-word response latency must be greater than word response latency, in computer simulations and perhaps in human perception as well.

Results from the three models were compared against actual data from a recent, purely experimental study of the lateralized lexical decision task (Weems & Zaidel, 2003). In that study, fifty participants were tested, all strongly right handed with no history of neurological illness. Two hundred eighty-eight words and non-words were used, each between three and five letters in length. In a manipulation not examined here, half of the words were high frequency (>100 instances per million per Francis and Kucera, 1982) and half were low frequency (<10 instances per million). All non-words were orthographically legal, matched with the words in length. Appendix B presents a list of the behavioral stimuli used. Each stimulus was shown twice, once in each visual field, for a total of 576 trials. Stimuli remained on a computer display screen for a total of 165 milliseconds, one degree of visual angle away from central fixation. Subjects responded bimanually by depressing buttons with both their left and right hands if the target was a word, and different buttons with both hands if the target was a non-word. Rest periods every 144 trials were provided, and testing lasted approximately 45 minutes. Both accuracy (percent errors) and latency (time between stimulus presentation and response) measures were collected, and only response latencies for correct trials were analyzed. Behavioral data from this study was subjected to a two-way, within subjects Analysis of Variance (ANOVA) with factors Visual Field and Target Wordness. A more detailed description of the experimental testing is described elsewhere (Weems & Zaidel, 2003).

RESULTS

Behavioral data. Figure 2 shows accuracy and latency results from each of the three computational models, as well as for the behavioral data. For accuracy, main effects of target wordness ($F_{1,49}$=10.64, p<.005) and visual field ($F_{1,49}$=67.53, p<.001) revealed that non-words were identified with fewer errors than words (10.3% and 14.3%) and that stimuli were identified more accurately in the RVF compared to the LVF (9.1% and 15.5%). For latency the same main effects were observed, as RVF trials were responded to more quickly than LVF trials (667.6 and 686.9 msec; $F_{1,49}$=17.08, p<.001), although words were identified earlier than non-words (665.2 and 689.3 msec; $F_{1,49}$=14.49, p<.001). These two factors interacted for both accuracy ($F_{1,49}$=23.82, p<.001) and latency ($F_{1,49}$=14.84, p<.001). Although both words and non-words were identified more accurately in the RVF ($F_{1,49}$=52.42, p<.001 and $F_{1,49}$=13.48, p<.001, respectively), the effect was much stronger for words, leading to a non-word advantage for LVF trials ($F_{1,49}$=19.28, p<.001) but not RVF trials. For latency, a RVF advantage was observed for words ($F_{1,49}$=23.44, p<.001) but not non-words, leading to a word advantage only for RVF trials ($F_{1,49}$=26.84, p<.001).

These results represent standard lexical decision findings, although the word advantage in response latency but not response accuracy is surprising. Although the word advantage in response latency is a reliable and robust finding for lateralized lexical decision, the word

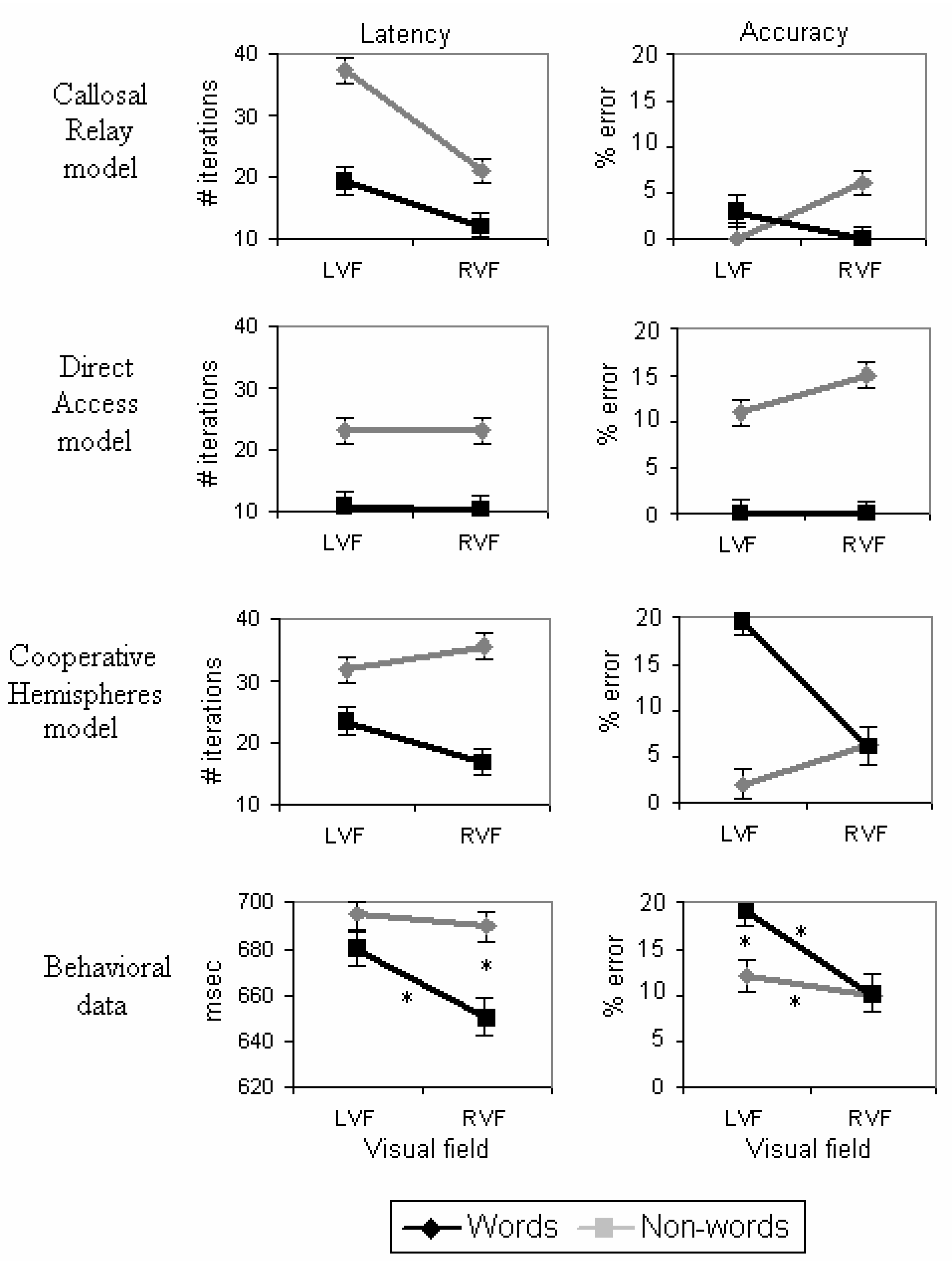

Figure 2. Results obtained from each of the three model types, as well as from behavioral data obtained separately. Vertical axes represent number of iterations before a response is made (models) or response latency (behavioral data) and percent incorrect identifications (accuracy; words identified as non-words and non-words identified as words). Error bars represent the standard error, and asterisks indicate significant effects in the behavioral results. While a RVF advantage occurs for word recognition only for both behavioral accuracy and latency, it occurs for the direct access and cooperative hemispheres models in response latency and for the callosal relay and cooperative hemispheres models in response accuracy.

advantage for response accuracy is common but less consistent (e.g., Faust, Kravetz, & Babkoff, 1993; Lambert, 1991). It is difficult to say why some studies show a word advantage for response accuracy while others do not. As described above regarding the neural models, it was impractical to set parameters such that non-word recognition was faster than word recognition, perhaps providing useful insight into why this also fails to occur in human recogntion. However, there was no such straight-forward relationship between word and non-word response accuracy, indicating that such measures may vary by subject to subject (or even experiment by experiment).

Callosal Relay model. For latency, both words and non-words showed a strong RVF advantage, with words being identified faster than non-words. Although a weak RVF advantage was observed for word recognition with response accuracy, this was overshadowed by a ceiling effect of performance and few errors overall. To explore the possibility that overlearning led hemispheric asymmetries to be washed out (i.e. that the choice of training epochs – 1000 – led the callosal relay model to show minimal asymmetries due to overlearning), performance of the model at shorter training intervals was explored. Results showed that performance between the visual fields was nearly identical even as early as 400 training epochs, with near perfect performance (<10% errors). In contrast to words, non-words showed a LVF advantage in response accuracy.

The lack of a strong RVF accuracy advantage for words, and the strong RVF latency advantage for non-words, suggests that performance of the callosal relay model differed significantly from the behavioral results. The hallmark sign of a left hemisphere advantage during lexical decision, namely a strong RVF advantage for word recognition, was not apparent in either dependent measure. It was surprising that such a strong RVF advantage should occur in non-word response latency, since these stimuli had not yet been presented to the model. Indeed, one might make the exact opposite prediction: it should take longer for activity to settle for the RVF (left hemisphere) due to the more strongly established (through the learning rate advantage) attractors generated by its weight matrix.

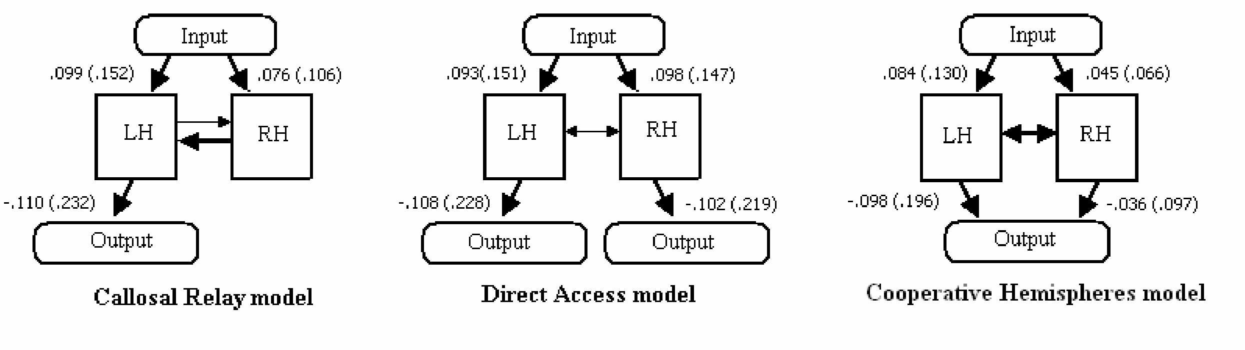

Figure 3. Mean connection weight values for each model type. Values in parentheses represent mean absolute weight values. Nearly identical levels of activity are sent to the left and right hemispheres for the direct access model. The callosal relay and cooperative hemispheres models, however, show a strong left hemisphere advantage in activity propagation (larger weight values).

Although it is a rather general measure that fails to capture much of the connectivity reflected in the neural network, mean connection weights between each hemisphere and the input/output layers were used to learn more about how each hemisphere contributed to responses. Figure 3 shows that connection weights between the input layer and the left and right hemisphere strongly differed across the three models. For the callosal relay model most activity was propagated to the left hemisphere rather than the right. This difference is observed for both mean connection weights, as well as the mean absolute value of those connection weights, which indicates that the left hemisphere received both stronger excitatory and inhibitory connections from the input layer. Training appears to have led to a minimization of right hemisphere involvement, with most activity (and stimulus processing responsibility) going to the left hemisphere.

One common experimental finding in behavioral lexical decision is that words that are more “word-like”, like those that have higher bigram frequency, are identified quicker and more accurately than those with fewer neighbors (see Westbury and Buchanan, 2002 for a review). In the current study, we also examined the effect of bigram frequency on word response latency by measuring the number of instances each letter pair occurs in the word corpus and examining the correlation between bigram frequency for each word and the number of iteration cycles before recognition. We chose to examine bigram frequency for words, rather than non-words, based on research showing that bigram frequency has the strongest effect for words, with little impact on non-word recognition (Rice & Robinson, 1975). To calculate bigram frequency, we added the number of instances that each letter pair occurs in the 60 word corpus, and calculate the mean bigram frequency for each word. We chose to examine the correlation between bigram frequency and response latency, rather than simply comparing bigram frequency for correctly and incorrectly identified words, because accuracy was so high that there were too few incorrectly identified words to make an accurate comparison. This relationship between “word-like-ness” and response latency may be difficult to assess; because so few letters were available, there was not the variance of bigram frequency in the current corpus that is encountered in the actual English language. For the callosal relay model, the correlation between word bigram frequency and response latency was -0.03. Thus, for this model bigram frequency failed to predict speed of word recognition.

Direct Access model. It is clear from Figure 2 that neither accuracy or latency performance demonstrated a RVF advantage. In fact, the only apparent visual field difference was a LVF advantage for non-word accuracy, mimicking the same finding from the callosal relay model. A word advantage was observed for both response accuracy and latency.

As with the callosal relay model, we also examined the possibility that the lack of visual field differences may be due to an overall ceiling effect for word recognition. However, like the callosal relay model, even as early as after 400 training epochs word recognition performance was near perfect (5-10% errors). It appears that when interaction between the hemispheres is minimized, each hemisphere becomes a competent, independent word recognizer. From Figure 3, we see that both hemispheres received strong connectivity with the input layer, even though the left was specialized for the task. This result is in contrast to the connectivity pattern with the callosal relay model, in which activity propagated substantially more to the left hemisphere. Connection weights to the respective output layers were also similar between the hemispheres, even though post-hoc analysis showed that the hemisphere receiving the primary input (left hemisphere for RVF and right hemisphere for LVF) was the one to respond first in all cases.

An examination of the relationship between bigram frequency, as calculated above, and response latency showed a weak, negative correlation: -0.13. The correlation did act in the anticipated direction, in that higher bigram frequency led to faster responses (fewer iterations to recognition). The modest degree of the relationship may be due to the inherent limitations from defining bigram frequency in the current situation, as discussed previously.

Cooperative hemispheres model. Unlike the callosal relay and direct access models, a RVF advantage was observed for accuracy of word recognition. Also of note is that this RVF advantage was limited to words, as non-words showed a weak LVF advantage for both dependent measures. Thus, the cooperative hemispheres model was the only model to mirror behavioral results in both response accuracy and latency, demonstrating a RVF advantage for words but not non-words

Like the callosal relay model, the right hemisphere was not strongly connected to the input layer (Figure 3), and was also more weakly connected to the output layer. Indeed, the input layer shared twice the connectivity with the left hemisphere than the right, even though the cooperative hemispheres model was the only one to allow free activity sharing between the hemispheres. Surprisingly, even the callosal relay model shared stronger connectivity between the input layer and right hemisphere, despite the fact that the right hemisphere only contributed indirectly (through the left hemisphere) towards response.

Finally, as with the direct access model, a modest negative correlation was observed between bigram frequency and word response latency (-0.18). Although the relationship was not strong, it did act in the anticipated direction: greater word-like-ness led to faster response.

DISCUSSION

As with many computational models, each of those discussed here is a gross simplification from actual cognitive strategies used during human performance, and the limited corpus of words used also limits the generality of our model. For example, as the stimuli were introduced to the hemispheres at the graphemic level without lower level analysis, and also presented without any context, this task surely differs from actual performance in many ways. However, these simplifications do not substantially diminish the significance of the observed results, since each is considered only relative to the others, and since it is how the hemispheres interact that was of primary interest. Indeed, it is in such simplicity that these models become their most informative; if interesting hemispheric asymmetries emerge even when lower level aspects of the stimulus processing are ignored, then such models may be useful in describing a variety of higher level cognitive tasks.

Both the cooperative hemispheres and callosal relay models showed some hemispheric specialization in performance accuracy, with the strongest hemispheric differences being in word response latency. These two models included hemispheric transfer (unlike the direct access model), thereby allowing activity to remain coordinated between the hemispheres during training and subsequent processing. It is notable that training involved bilateral presentation, so the two hemispheres were trained to work together rather than individually. As a result, any hemispheric specialization emerged from the simple dynamics defined by the models, rather than any rules imposed by the model design or needs of the task. This specialization was not dependent on asymmetric transfer between the hemispheres, as it occurred both for the cooperative hemispheres model (in which activity traveled freely in both directions between the hemispheres), and the callosal relay model (in which transfer was primarily right to left).

Mean connection weight values between the primary and association layers (Figure 3) showed that the callosal relay and cooperative hemispheres models relied much less on the right hemisphere for response. This laterality was observed in contrast to the direct access model, in which strong connectivity emerged to both hemispheres and word recognition was high for both visual fields (although faster for the RVF). The lateralization of the cooperative hemispheres and callosal relay models led to different outcomes: the callosal relay model showed faster recognition for both words and non-words for RVF trials, whereas the cooperative hemispheres showed a RVF response latency advantage only for words. Thus, the cooperative hemispheres model was the only one to mimic behavioral results in both accuracy and latency performance (i.e., the only model that showed a RVF advantage in both response measures for words only).

The strong asymmetry observed when interhemispheric connectivity was included runs counter to some neurological evidence that shows greater interhemispheric connectivity is associated with reduced asymmetry. For example, Aboitiz et al (1992) observed using post-mortem analysis that greater hemispheric asymmetry in perisylvian size in males is related to reduced isthmus size. Similar results were observed using magnetic resonance imaging, in which greater hemispheric asymmetry for males was associated with lower corpus callosum size (Dorion et al., 2000). The difference between these anatomic findings (although still somewhat controversial) and our neural network results may lie in the different metrics used for asymmetry. Anatomically, asymmetry is measured using cortical size, whereas here asymmetry is defined functionally. Perhaps some interhemispheric communication is necessary for left hemisphere specialization to occur because, without such communication, each hemisphere learns independently (as discussed next regarding the direct access model). In the brain, where the two hemispheres clearly work in tandem, the two hemispheres do not get the opportunity to learn independently, and so increased inter-hemispheric connectivity can have limiting effects on asymmetry.

In contrast, the successful performance of the direct access model, regardless to which visual field the stimuli were presented, suggests that this model allowed each hemisphere to become a strong word recognizer. Bilateral competence was supported by the fact that the hemisphere receiving the bulk of stimulus information was the one making responses, with the left hemisphere being slightly faster for identifying words. However, the strong recognition ability of both hemispheres in this model runs counter to the behavioral evidence presented here and elsewhere (Bradshaw & Gates, 1978; Chiarello, 1985; Leiber, 1976). For example, word recognition ability of the disconnected right hemisphere, a measure only able to be achieved with split-brain patients, is typically very low or non-existent (Zaidel, 1990, 1998). For this reason, it remains unlikely that the direct access model, at least as defined here, is a completely adequate model for hemispheric word recognition.

It is surprising that the model that included symmetric transfer between the hemispheres (cooperative hemispheres model), not asymmetric transfer (callosal relay), was the only one to show hemispheric specialization effects for both response measures. The cooperative hemispheric model is special because it makes no explicit assumptions about which hemisphere is taked with processing a stimulus or making a response. Indeed, it is difficult to identify which hemisphere holds the greatest responsibility for making a response because both are connected to the same output mechanisms; any output is based on the summed output of both hemispheres. The cooperative hemispheres model, therefore, was the only one to allow intrinsic differences between the hemispheres (in this case, learning rate) to govern asymmetries, rather than constraints of the model itself.

Finally, the relationship between bigram frequency and response latency showed that, for the cooperative hemispheric and direct access models, more word-like words are identified faster than those that are less word-like. Here, being word-like is defined as sharing bigrams that also commonly co-occur in other lexical entries. However, this correlation was not exceptionally strong, reaching -.018 and –0.13, respectively. Perhaps with a larger vocabulary the correlation would have been greater, as 60 words provided only a minimal chance to examine correlations. A wider variance of bigram frequency might also contribute to greater power to examine such relationships; having only 15 letters led to few bigrams and therefore little variance in bigram frequency. Because stimuli were resented in a 3x5 input matrix, rather than serially or some more realistic manner, the model also had no way of assessing important orthographic factors such as letter position. Thus, bigram frequency is a relatively artificial construct in the current model, although still worthy of inspection. However, the correct direction of the trend in these two models is promising. The lack of a similar significant correlation in the direct access model might be linked with the overall fast word recognition it exhibited (i.e., minimal variance in response latency), although further examination is warranted.

APPENDIX A

Input matrix

Row 1 node 1:B, node 2:H, node 3:L, node 4:P, node 5:W
Row 2 node 1:A, node 2:E, node 3:I, node 4:O, node 5:U
Row 3 node 1:D, node 2:G, node 3:M, node 4:N, node 5:T

Words

bad
bag
ban
bat
bed
beg
bet
bid
big
bin
bit
bog
bud
bug
bum
bun
but
had
ham
hat
hem
hen
hid
him
hit
hog
hot
hug
hum
hut
lad
lag
led
leg
let
lid
lit
log
lot
lug
pad
pan
pat
peg
pen
pet
pig
pin
pit
pod
pot
pun
put
wag
wed
wet
wig
win
wit
won

Non-words

bam
bem
ben
bim
bom
bon
bot
han
hed
heg
het
hig
hin
hod
hom
hon
hud
lam
lan
lat
lem
len
lig
lim
lin
lod
lom
lon
lud
lum
lun
lut
pag
ped
pem
pid
pim
pog
pom
pon
pud
pum
wad
wam
wan
wat
weg
wem
wen
wid
wim
wod
wog
wom
wot
wud
wug
wum
wun
wut

APPENDIX B

Words – behavioral data

able
ache
act
acute
age
agile
aisle
alone
amuse
angel
ant
apron
army
art
ask
ate
attic
avid
awe
bad
ball
bed
been
beg
bet
big
bike
bin
bind
bit
bite
black
blend
blood
blush
body
boot
boy
brisk
brood
brown
bud
camel
canon
car
care
chalk
chant
chasm
chill
cigar
class
clear
close
clue
color
comb
come
comic
crisp
crown
cub
curl
cut
dash
date
day
dig
dim
does
done
door
doubt
drive
drum
due
duel
dwarf
early
egg
ego
elm
end
epic
few
final
find
fix
flood
floor
found
fox
frost
fry
game
gap
gem
girl
give
glove
gown
great
grill
gun
hair
halt
hand
has
hawk
heard
heavy
heir
help
herb
high
hike
hip
hit
hold
hook
hope
horse
hot
house
hymn
idea
kind
kite
know
law
lay
lazy
lead
leap
leash
leave
led
let
lick
lid
life
light
local
logic
lost
lot
love
low
major
make
man
mar
mass
might
mild
mole
money
moral
move
mow
mug
music
nail
name
noisy
oak
off
old
party
pay
peace
peel
pill
plant
point
pool
press
pry
put
quart
quest
race
rage
ram
ran
rat
reach
ready
real
red
rhyme
rid
rinse
road
robin
rot
rub
rude
sail
sat
say
scare
scene
see
set
sew
shine
short
shy
ski
soup
spa
spell
stage
stain
stand
stop
straw
sue
sun
sword
tag
take
tap
tax
tin
top
total
tow
truce
truth
try
turn
twin
type
unite
use
vent
view
voice
wall
want
war
wasp
water
way
week
wife
wig
wink
woman
word

bump
buoy
burn
bush
calf
call
erupt
eye
face
fade
fairy
far
ink
issue
ivy
jam
job
kick
order
oven
owe
own
pan
paper
skip
sly
small
sniff
snore
soak
world
yacht
yea
yearn
yet
zoo

Non-words – behavioral data

abes
aby
aff
aften
aib
aket
ank
aprit
arl
awb
balp
bim
bir
biue
boj
booce
browl
bworl
caoul
cealp
ceasp
ced
cew
chae
chig
chrea
cipe
clell
cobin
crael
crean
crice
crill
cumis
cur
danh
dayv
duby
dush
dwo
dyb
dystu
eakem
ect
edis
eefu
efec
eftom
ega
elif
emd
eme
emin
enco
enole
enot
epho
erak
essek
esu
eusac
eveol
evom
ewa
fambi
fas
faw
feli
feri
fiech
flude
fofto
foo
foten
giwn
gnith
grue
gubda
gwe
huc
ind
inl
intle
ipy
jelp
jind
jorl
jov
jugh
kaap
kade
kalem
kard
karm
karsp
kere
ket
kewe
kide
kleap
klim
klo
knue
koes
kos
krasp
kug
kuy
kworp
lacle
laipt
malk
marf
mimk
moce
moge
molp
myro
neop
nihc
nowgy
nus
odle
odolb
oftes
ouy
ovep
owk
phoag
pih
plock
poj
prisc
raih
rairt
rajom
reaj
ree
reoy
resu
rieh
rik
riker
rimic
rine
rolam
rolic
romo
shurd
skoob
skup
sleed
sneav
snep
sni
soms
sorhe
spack
spo
spon
sreps
srill
stamm
stass
stin
stip
stoce
stol
stosp
susei
talpo
tatur
tec
tesb
theex
tift
tinop
tivis
tiwch
tof
tog
toh
tosp
tov
tra
uke
undof
unerg
urt
ute
vacht
vad
vadsi
vige
vight
vired
voght
vole
vor
vox
vra
vro
vug
vyheg
walp
wapte
wapz
welob
wevi
witer
wols
wom
wug
xis
yad
yag
yare
yas
yatol
yats
yax
yed

deaty
deha
dem
dhol
dinf
diy
doi
dorw
drelo
drila
droor
fov
foz
fult
fux
galk
gax
geami
geb
gep
ger
gip
lalc
lar
lare
laz
lifun
lin
lirg
liw
lorch
lum
lund
rons
ruof
rusic
ryc
sacla
saj
sar
scyne
sdu
seab
sert
treal
tren
trosh
trown
tsik
turp
tus
tuz
twip
tworc
tyn
yelc
yis
yoc
yoom
yow
yub
zay
zimat
zix
zow
zwap

REFERENCES

Aboitiz, F., Scheibel, A. B., Fisher, R., & Zaidel, E. (1990). Individual differences in brain asymmetries and fiber composition in the human corpus callosum. *Brain Research Reviews, 598*(1-2), 154-161.

Alvarez-Linera, J., Martin-Plasencia, P., Maestu-Unturbe, F., Sola, R., Iglesias, R., & Serrano, J. (2002). Hemispheric dominance for language and functional magnetic resonance: A comparison of three tasks. *Review of Neurology, 35*(2), 115-118.

Anninos, P., Argyrakis, P., & Skouras, A. (1984). A computer model for learning processes and the role of the cerebral commissures. *Biological Cybernetics, 50*, 329-336.

Anninos, P., & Cook, N. D. (1988). Neural net simulation of the corpus callosum. *International Journal of Neuroscience, 38*, 381-391.

Babkoff, H., & Ben-Uriah, Y. (1983). Lexical decision time as a function of visual field and stimulus probability. *Cortex, 19*, 13-30.

Banich, M., & Federmeier, K. (1999). Categorical and metric spatial processes distinguished by task demands. *Journal of Cognitive Neuroscience, 11*(2), 153-166.

Binder, J., Frost, J., Hammeke, T., Cox, R., Rao, S., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience, 17*(1), 353-362.

Bogen, J. (2000). Split-brain basics: relevance for the concept of one's other mind. *Journal of American Academy of Psychoanalysis, 28*(2), 341-369.

Bradshaw, J., & Gates, E. (1978). Visual field differences in verbal tasks. Effects of task familiarity and sex of subject. *Brain and Language, 5*, 166-187.

Calandra-Buonaura, G., Basso, G., Gorno-Tempini, M., Serafini, M., Pagnoni, G., Baraldi, P., Porro, C., & Nichelli, P. (2002). Human brain language processing areas identified by functional magnetic resonance imaging using a lexical decision task. *Functional Neurology, 17*(4), 183-191.

Caplan, D., Hildebrandt, N., & Makris, N. (1996). Localization of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain, 119*, 933-949.

Chiarello, C. (1985). Hemisphere dynamics in lexical access: automatic and controlled priming. *Brain and Language, 26*, 146-172.

Chiarello, C. (1988). Lateralization of lexical processes in the normal brain: A review of visual half-field research. In H. Whitaker (Ed.), *Contemporary reviews in neuropsychology* (pp. 36-76). New York: Springer-Verlag.

Christman, S., & Hackworth, M. (1993). Equivalent perceptual asymmetries for free viewing of positive and negative emotional expressions in chimeric faces. *Neuropsychologia, 31*, 621-624.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review, 100*(4), 589-608.

Coney, J. (2002). The effect of associative strength on priming in the cerebral hemispheres. *Brain and Cognition, 50*, 234-241.

Cook, N. D. (1986). *The brain code: mechanisms of information transfer and the role of the corpus callosum*. New York: Menthuen Press.

Cook, N. D., & Beech, A. (1990). The cerebral hemispheres and bilateral neural nets. *International Journal of Neuroscience, 52*, 201-210.

Corsi-Caberra, M., Gutierrez, S., Ramos, J., & Arce, C. (1988). Interhemispheric correlation of EEG activity during successful and unsuccessful cognitive performance. *International Journal of Neuroscience, 39*, 253-259.

Dorion, A., Chantome, M., Hasboun, D., Zouaoui, A., Marsault, C., Capron, C., & Duyme, M. (2000). Hemispheric asymmetry and corupus callosum morphometry: a magnetic resonance imaging study. *Neuroscience Research, 36*(1), 9-13.

Ellis, A., & Young, A. (1988). *Human Cognitive Neuropsychology*. East Sussex, England: Psychology Press.

Ellis, A., Young, A., & Anderson, C. (1988). Modes of word recognition in the left and right cerebral hemispheres. *Brain and Language, 35*, 254-273.

Engel, A., Konig, P., Kreiter, A., & Singer, W. (1991). Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science, 25*, 1177-1179.

Faust, M., Kravetz, S., & Babkoff, H. (1993). Hemisphericity and top-down processing of language. *Brain and Language, 44*(1), 1-18.

Francis, W., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston, MA: Houghton Mifflin.

Grindrod, C., & Baum, S. (2003). Sensitivity to local sentence context information in lexical ambuguity resolution: evidence from left- and right-hemisphere damaged individuals. *Brain and Language, 85*(3), 503-523.

Grodzinski, Y. (2000). The neurology of syntax: language use without Broca's area. *Behavioral & Brain Sciences, 23*(1), 21-71.

Hagoort, P. (1998). The shadows of lexical meaning in patients with semantic impairments. In B. Stemmer & H. Whitaker (Eds.), *Handbook of Neurolinguistics* (pp. 235-248). San Diego, CA: Academic Press.

Houzel, J., Carvalo, M., & Lent, R. (2002). Interhemispheric connections between primary visual areas: beyond the midline rule. *Brazilian Journal of Medicine and Biological Research, 35*(12), 1441-1453.

Howard, N., & Reggia, J. (In press). The effects of multi-task learning and time-varying hemispheric asymmetry on lateralization in a neural network model. *Laterality*.

Hugdahl, K., & Davidson, R. (2003). *The asymmetrical brain*. Cambridge, MA: MIT Press.

Iacoboni, M., & Zaidel, E. (1996). Hemispheric independence in word recognition: evidence from unilateral and bilateral presentations. *Brain & Language, 53*(1), 121-140.

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

Knyazeva, M., Kiper, D., Vildavski, Y., Despland, A., Maeder-Ingvar, M., & Innocent, G. (1999). Visual stimulus-dependent changes in interhemispheric EEG coherence in humans. 3095-3107.

Lambert, A. (1991). Interhemispheric interaction in the split-brain. *Neuropsychologia, 29*(10), 941-948.

Leiber, L. (1976). Lexical decisions in the right and left cerebral hemispheres. *Brain and Language, 3*, 443-450.

Mima, T., Oluwatimilehin, T., Hiraoka, T., & Hallet, M. (2001). Transient interhemispheric neuronal synchrony correlated with object recognition. *The Journal of Neuroscience, 21*(11), 3942-3948.

Mohr, B., Pulvermuller, F., & Zaidel, E. (1994). Lexical decision after left, right and bilateral presentation of function words, content words and non-words: evidence for interhemispheric interaction. *Neuropsychologia, 32*(1), 105-124.

Pulvermueller, F., Harle, M., & Hummel, F. (2001). Walking or talking? Behavioral and neurophysiological correlates of action verb processing. *Brain and Language, 78*, 143-168.

Pulvermuller, F., Haerle, M., & Hummel, F. (2000). Neurophysiological distinction of verb categories. *Neuroreport: For Rapid Communication of Neuroscience Research, 11*(12), 2789-2793.

Pulvermuller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex, 9*(5), 497-506.

Pulvermuller, F., & Mohr, B. (1996). The concept of transcortical cell assemblies: A key to the understanding of cortical lateralization and interhemispheric interaction. *Neuroscience & Biobehavioral Reviews, 20*(4), 557-566.

Reggia, J., Goodall, S., & Shkuro, Y. (1998). Computational studies of lateralization of phoneme sequence generation. *Neural Computation, 10*(5), 1277-1297.

Reggia, J., & Levitan, S. (2003). Hemispheric interaction and specialization. In M. Arbib (Ed.), *Handbook of Brain theory and Neural Networks* (Second ed., pp. 525-528). Cambridge, MA: MIT Press.

Reinhard, J., & Trauzettel-Klosinski, S. (2003). Nasotemporal overlap of retinal ganglion cells in humans: a functional study. *Investigatory Opthalmology, 44*(4), 1568-1572.

Rice, G., & Robinson, D. (1975). The role of bigram frequency in the perception of words and non-words. *Memory and Cognition, 3*(5), 513-518.

Ringo, J., Doty, R., Demeter, S., & Simard, P. (1994). Time is of the essence: A conjecture that hemispheric specialization arises from interhemispheric conduction delay. *Cerebral Cortex, 4*, 331-343.

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

Seidenberg, M., Plaut, D., Petersen, A., McClelland, J., & McRae, K. (1994). Nonword pronunciation and modelsof word recognition. *Journal of Experimental Psychology: Human Perception & Performance, 20*(6), 1177-1196.

Shevtsova, N., & Reggia, J. (1999). A neural network model of lateralization during letter identification. *Journal of Cognitive Neuroscience, 11*(2), 167-181.

Shillcock, R., Ellsion, T., & Monaghan, P. (2000). Eye-fixation behavior, lexical storage, and visual word recognition in a split processing model. *Psychological Review, 107*(4), 824-851.

Shkuro, Y., Glezer, M., & Reggia, J. (2000). Interhemispheric effects of simulated lesions in a neural model of single-word reading. *Brain and Language, 72*, 343-374.

Tucker, D. (1987). Hemisphere specialization: A mechanism for unifying anterior and posterior brain regions. In D. Ottoson (Ed.), *Duality and unity of brain* (pp. 180-193). London: MacMillan Press.

Tucker, D., & Williamson, P. (1984). Asymmetric neural control systems in human self-regulation. *Psychological Review, 91*, 185-215.

Weems, S., & Zaidel, E. (2003). Repetition priming in the two cerebral hemispheres for attended and unattended lexical stimuli. *Submitted.*

Weiss, S., & Mueller, H. (2003). The contribution of EEG coherence to the investigation of language. *Brain and Language, 85*(2), 325-343.

Westbury, C., & Buchanan, L. (2002). The probability of the least likely non-lentgh-controlled bigram affects lexical deicision react.

Young, A., & Ellis, A. (1985). Different methods of lexical access for words presented in the left and right visual fields. *Brain and Language, 24*, 326-358.

Zaidel, E. (1979). On measuring hemispheric specialization in man. In B. Rybak (Ed.), *Advanced Technobiology* (pp. 365-404). Paris: Sijthoff and Noordhoff, Alphen aan den Rejn.

Zaidel, E. (1990). Language functions in the two hemispheres following complete cerebral commissurotomy and hemispherectomy. In R. D. Nebes & S. Corkin (Eds.), *Handbook of neuropsychology, Vol. 4.* (pp. 115-150). New York, NY, US: Elsevier Science.

Zaidel, E. (1998). Language in the right hemisphere following callosal disconnection. In B. Stemmer & H. Whitaker (Eds.), *Handbook of Neurolinguistics* (pp. 369-383). San Diego, CA: Academic Press.

Zaidel, E., Clarke, J. M., & Suyenobu, B. (1990). Hemispheric independence: A paradigm case for cognitive neuroscience. In A. B. Scheibel & A. F. Wechsler (Eds.), *Neurobiology of Higher Cognitive Function.* (pp. 297-355). New York, NY, US: The Guilford Press.

Zouridakis, G., Simos, P., Breir, J., & Papanicolaou, A. (1998). Functional hemispheric asymmetry asessment ina visual language task using MEG. *Brain Topography, 11*(1), 57-65.

Scott A. Weems, Departments of Computer Science and Neurology; James A. Reggia, Departments of Computer Science and Neurology.

This investigation was supported by NIH awards T32 DC00061 and RO1 NS35470..

Correspondence concerning this article should be addressed to Scott Weems, Department of Computer Science, University of Maryland at College Park, A.V. Williams Building, College Park, Maryland 20742. Phone: (301) 405-2670. Fax: (301) 405-6707. Email: sweems@cs.umd.edu.