Two Hemispheres – Two Networks: A Computational Model Explaining Hemispheric Asymmetries While Reading Ambiguous Words

Orna Peleg[1], Larry Manevitz[2], Hananel Hazan[2] and Zohar Eviatar[1]

Institute of Information Processing and Decision Making[1]
and
Department of Computer Science[2]
University of Haifa, Israel
{*opeleg,zohare*}*@research.haifa.ac.il,*
{*manevitz, hhazan01*}*@cs.haifa.ac.il*

March 9, 2008

Abstract

A computational model for reading that takes into account the different processing abilities of the two cerebral hemispheres is presented. This dual hemispheric reading model follows the computational lines of Seidenberg and McClelland [41], and especially Kawamoto [23]; but postulates a difference in architecture between the right and left hemispheres. Specifically it is assumed that orthographic, phonological and semantic units are completely connected in the left hemisphere, while there are no direct connections between phonological and orthographic units in the right hemisphere. It is claimed that this architectural difference results in hemisphere asymmetries in resolving lexical ambiguity and more broadly in the processing of written words.

Simulation results bear this out. First, we show that the two networks successfully simulate the time course of lexical selection in the two cerebral hemispheres. Further, we were able to see a computational advantage of two separate networks, when information is transferred from the right hemisphere network to the left hemisphere network. Finally, beyond reproducing known empirical data, this dual hemispheric reading model makes novel and surprising predictions that were found to be consistent with new human data

1 Introduction

Human understanding of written words requires accessing and integrating different sources of information from long-term memory. This process is complicated by the fact that many words have more than one distinct meaning (e.g., the homograph *bank* is associated with a financial institution or a riverside), and thus part of the comprehension process entails selection of one of these meanings.

How do readers resolve lexical ambiguity? Ample evidence from behavioral research indicates that this selection process is governed by lexical and contextual factors. First, a particular meaning of a homograph may be more frequent or dominant than another. Second, the particular context in which the homograph is embedded may be biased toward one particular interpretation (e.g., Duffy, Morris & Rayner [9]; Titone [47]; Peleg, Giora & Fein [35, 36], Peleg & Eviatar [33]).

Although effects on ambiguity resolution are still debated (for an overview, see Simpson [44], [43]; Small, Cottrell, & Tanenhaus [45]), the majority of the semantic-priming literature suggests that when readers encounter an ambiguous word, all meanings become activated initially. However, following this brief exhaustive access stage, contextual and lexical factors lead to a selection of one particular meaning by enhancing activation of frequent and/or contextually-relevant meanings while at the same time suppressing activation of less frequent and/or contextually irrelevant meanings. This time course of lexical selection was successfully simulated by Kawamoto's simple recurrent network described below.

1.1 A Connectionist Approach to Lexical Ambiguity Resolution

A connectionist account of lexical ambiguity resolution was presented by Kawamoto [23]. In his fully recurrent network, ambiguous and unambiguous words are represented as a distributed pattern of activity over a set of simple processing units. More specifically, each lexical entry is represented over a 216 - bit vector divided into separate sub-vectors representing the "spelling", "pronunciation", "part of speech" and "meaning". The network is trained with a simple error correction algorithm by presenting it with the pattern to be learned. The result is that these patterns (the entire word including its orthographic, phonological and semantic features) become "attractors" in the 216-dimensional representational space (Hopfield [21]). The network can then be tested by presenting it with just part of the lexical entry (e.g., its orthographic pattern) and measuring how long various parts of the network take to settle into a pattern corresponding to a particular lexical entry. Kawamoto trained his network in such a way that the more frequent combination for a particular orthographic representation was the "deeper" attractor; i.e. the completion of the other features (semantic and phonological) would usually fall into this attractor. (This was accomplished by biasing the learning process of the network.). However, using a technological analogy of "priming" to bias the appropriate completion, the resulting attractor could in fact be the less frequent combination - which corresponds nicely to human behavioral data. Indeed, consistent with human empirical results, after the network was trained, the resolution process was affected by the frequency of the different lexical entries (reflected in the strength of the connections in the network) and by the context

1.2 Ambiguity resolution in the Two Hemispheres The Standard Model

Hemispheric asymmetries were found to be of particular importance in the processing of ambiguous words because both context and frequency have been shown to have differential implications for the processing of language in the hemispheres (e.g., Beeman, Friedman, Grafman, Perez, Diamond & Lindsay [2]; Faust & Gernsbacher [13]; Peleg & Eviatar [33]). Moreover, these studies show that the process of ambiguity resolution requires the intact functioning of both cerebral hemispheres (e.g., Grindrod, Baum and Shari [18]; Mason & Just [29]).

Importantly, several studies (e.g., Burgess & Simpson [6]; Faust & Gernsbacher [13]; Faust & Chiarello [12]) have shown that the time course of lexical selection may be different for the left than for the right hemisphere. According to these studies, the left hemisphere (LH) quickly selects one meaning (the contextually compatible meaning when prior contextual information is biased, or the salient, more frequent meaning when embedded in non-constraining contexts), whereas the right hemisphere (RH) maintains alternative meanings (including less salient, subordinate and contextually inappropriate meanings). In the literature, this proposal is referred to as the "standard model" of hemispheric differences in meaning resolution.

Four major proposals have been advanced to account for the sustained activation of less frequent andor contextually incompatible meanings in the RH as opposed to their fast decay in the LH. First, according to The "Coarse Coding Model" suggested by Beeman [4], [3], meaning representations in the LH are finely-coded (narrow representations that include only closely related meanings), whereas semantic representations in the RH are coarsely coded (broader representations that include less-related meanings as well). In addition, several researchers proposed that hemispheric differences in word meaning activation result from a selection mechanism, specific to LH processing, that inhibits or suppresses less related meanings (e.g., Tompkins [48]). Another explanation is that the RH is less sensitive to sentence-level information (Faust [11]). As a result, sentential information cannot be used for selection.

Finally, Burgess and Lund [5] suggested that differences in speed of activation onset could account for differences in meaning activation. In this view, meaning dominance lead to both stronger and longer activations of word meanings for both LH and RH processing. As a result, less-related meanings decay faster. However, because RH processing has a slower onset of speed activation, less related meanings are still activated at a point where they are already suppressed in the LH. In the following, we present an alternative explanation for hemisphere asymmetries in ambiguity resolution. In our proposal, meaning activation discrepancy is obtained without needing to postulate variant onset or decay speeds for the hemispheres; rather it is a consequence of the architectural choices of the hemispheric reading networks.

1.3 An Alternative Proposal

We suggest an alternative explanation for the observed hemisphere asymmetries in resolving lexical ambiguity. Our explanation relates to the different ways in which meanings are accessed in the two hemispheres; Generally speaking, there are two ways to access meaning from print: The visual route (from orthography directly to meaning), and the phonological route (from orthography to phonology to meaning). The visual route is believed to exist in both hemispheres. The phonological route, however, is available only to the left hemisphere (Zaidel [52], [53], [54]; Iacoboni & Zaidel, [22]). In principle two are better than one; since in the LH words can be read both visually and phonologically it is usually the faster and more accurate hemisphere.

We propose a simple model (see Figure 1) that incorporates a right hemisphere reading network and a left hemisphere reading network that differ in the coordination and relationships between orthographic, phonological and semantic representations. As in the "triangle" model (Seidenberg and McClelland [41]; and see also, Plaut, McClelland, Seidenberg and Patterson [39]; Harm & Seidenberg [19], Thivierge, Titone & Shultz [46]), in the LH, orthographic, phonological and semantic codes are fully interconnected. Importantly, however, in the RH, orthographic and phonological codes are not directly connected.

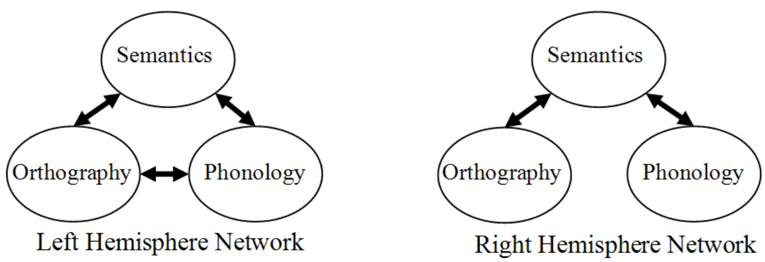

Figure 1: **The Dual Hemispheric Reading Model**

Specifically, in the LH, the orthographic representation of the word, automatically and directly activates both the phonological representation and the semantic representation of that word, whereas in the RH, orthography and phonology are not directly related, so that phonological representations, in the RH, are semantically mediated. The two structures are homogeneous in the sense that all computations involve the same sources of information. However, the time course of meaning activation and the relative influence of different sources of information at different points in time during this process is different, because these sources of information relate to each other in different ways.

To test this hypothesis, we implemented *two* artificial neural networks, [1] one for each hemisphere, and simulated the processing of two types of homographs: homophonic homographs (a single orthographic and phonological representation associated with multiple meanings, such as *bank*) and heterophonic homographs (a single orthographic representation associated with multiple phonological codes each associated with a different meaning, such as *bow*)

2 The Dual Hemispheric Model of Reading

The dual hemispheric model of reading is based on Kawamoto's [23] simple recurrent neural network presented above (see section 1.1). The model includes a LH network and a RH network (see Figure 2). There are 256 units in each network and each unit corresponds to one of the 256 features representing a lexical entry (*described below*). The LH network is a fully recurrent network: each unit receives input from the environment as well as from every other unit in the network. The RH network is identical to the LH, except that direct connections between units representing phonological features ("pronunciation" sub-vector) and units representing orthographic features ("spelling" sub-vector) were removed. Training and testing procedures were identical for both networks.

Forty eight patterns were created to represent 48 Hebrew (3-letter) words: 16 pairs (32 words) of homographs (both homophonic and heterophonic) and 16 unambiguous words. The homographs (e.g., *bank*) were all polarized, with one dominant meaning (e.g., "a financial institution") and one less frequent interpretation (e.g., "river side"). As a control, the unambiguous words were also divided into two groups: eight frequent words (as frequent as the dominant meaning of each homograph) and eight less-frequent words (to match the subordinate meaning of each homograph). Each lexical entry was represented

[1] A comment as to the role of computational models and simulations in such studies. Beyond the usual arguments that simulations force precision in theories, there is the additional fact that, because of the complexity, cognitive theories are always under-determined. (In other words, one can find competing explanations for the same data.) It is sometimes argued that such models should be as detailed as possible modeling, e.g. the internal physiological structure of the human. However, from the computational view, it is important to try to see if another instance of the theory, having the same capabilities that the theory posits also produces the computational results. From this outlook, it is actually the *simplest* model having this ability that gives the strongest support for the theory; and computational models are thus appropriate. For examples of computational work related to the subject of this paper, see Sejnowski and Rosenberg [42], Seidenberg and McClelland [41], Plaut [38], [37], Manevitz and Zemach [25], McClelland, Seidenberg and Patterson [39], Plaut and Shallice [40], Hinton and Shallice [20].

Of course, if the model makes additional predictions, which are borne out in human experimental data then this also strengthens the theory. Obviously, there is much room for interactions comparing results and designing experiments between both the psycho-physical and the computational experiments. In the issue under investigation here (the time-course of lexical selection in the two hemispheres), both computational and human experiments were performed. In this paper, our main focus is on the computational results and only the most significant results from the human experiments are mentioned. Full details on the human experiments and additional computational simulations will appear in Peleg, Eviatar, Hazan and Manevitz, in preparation.

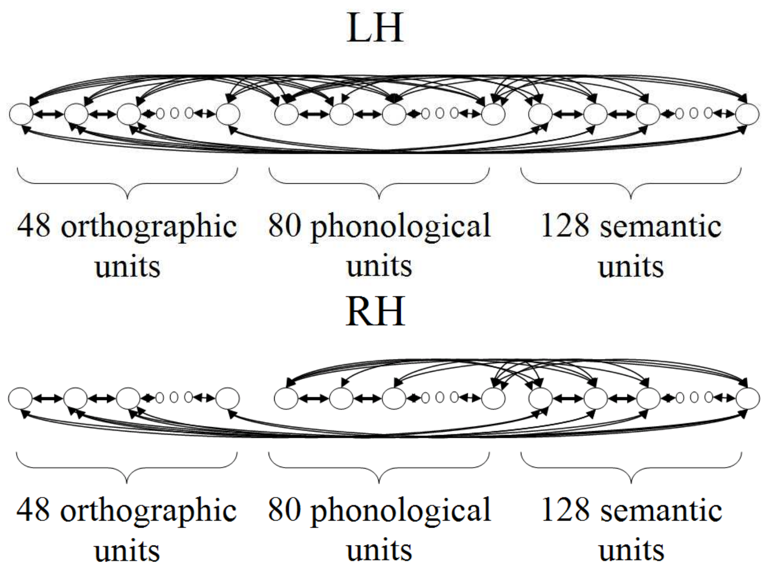

Figure 2: **Architecture of the two connectionist networks.** There are 256 units in each network representing the spelling (orthography), pronunciation (phonology) and meaning of words. In the LH network (top panel), all units are connected to each other. In the RH network (bottom panel), orthographic and phonological units are not directly connected

by 16 groups of features (Each group was represented by 16 bipolar [-1, 1] features): Three orthographic groups of features represented its spelling (one group for each letter); 5 phonological groups of features represented its pronunciation (one group for each phoneme)[2] and 8 semantic groups of features represented its meaning. Thus, for each entry, 48 features represented the word's spelling (orthographic sub-vector), 80 features represented its pronunciation (phonological sub-vector), and 128 features represented its meaning (semantic sub-vector). Overall, each entry is represented as a vector of 256 bipolar bits.

In the training stage, an entry is presented to the network. This activates the corresponding units in the network and sets the activation level to the appropriate value: +1 if the feature is present, or -1 if the feature is absent. For each unit, the net input from all the other units in the network, weighed by the connection strength from a unit, is computed. After each learning trial, the connection strengths are modified with a simple error correction algorithm:

$$\Delta W_{ij} = \eta(target_i - input_i)target_i,$$

where $input_i = \sum_j W_{ij} target_j$, η is a scalar learning constant fixed to 0.00003, $target_i$ and $target_j$ are the target activation levels of units i and j, and $input_i$

[2]Since vowels are mostly deleted in Hebrew orthography, a three letter word can actually represent five sounds. For example, the Hebrew word for book is spelled "sfr" and pronounced sefer

is the net input to unit i. The magnitude of the change in connection strength is determined by the magnitude of the learning constand and the magnitude of the error $(target_i - input_i)$. [3]

After the networks were trained and thus the values of the connection strength have been set, the networks were tested by presenting just the orthographic part of the entry as the input (to simulate neutral context) or by presenting part of the semantic sub-vector together with the orthography (to simulate contextual bias). In each simulation the input sets the initial activation of the units. The level was set to +0.25 if the corresponding input feature was positive, -0.25 if it was negative and 0 otherwise. The activity of a single unit in the network is represented as a real value ranging between -1.0 and +1.0. This activity is determined by the input from the environment, the units connected to it, and the decay in its current level of activity. These influences lead to changes in the activity of a unit as a function of time (where time changes in discrete steps). That is, the activity of a unit (a) at time t + 1 is:

$$a(t+1) = Limit[\delta a(t) + [\sum_j W_{ij}(t)a_j(t)] + s_i(t)]$$

where δ is a decay variable that changes from 0.6 to 1 as the iterations increase, $s_i(t)$ is the influence of the input stimulus on unit a_i at time $t+1$, and *Limit* is a function that bounds the activity to the range from -1.0 to +1.0. [4] (This organization is taken from Kawamoto [23].)

In order to assess lexical access, the number of iterations through the network for all the units in the spelling, pronunciation or meaning fields to become saturated was measured. A response was considered an error if the pattern of activity did not correspond with the expected completion of the input, or if all the units did not saturate after 50 iterations. Activation of dominant and subordinate meanings of a given homograph was also examined as a function of time.

3 Simulations

3.1 Step 1 - Simulating the time course of lexical selection in the two hemispheres when homographs are presented in isolation (without context)

Twelve pairs of LH and RH networks were used to simulate 12 subjects in an experiment. The networks in each hemisphere differed on their randomly chosen initial connections weights (chosen within the range $-\eta$ to $+\eta$) and on the random order in which the words were presented. In all other respects, the networks were identical. On each learning trial an entry was selected randomly

[3] The small learning constant was found to be necessary because of the need to establish separate attractors for the two meanings of the homophones, which have a relative small hamming distance between them.

[4] The δ term was needed to avoid local minima.

		LH network	RH network
Homograph	Dominant/Frequent meaning	11.59	21.35
	Subordinate/less-frequent meaning	-	-
Unambiguous word	Frequent word	9.29	10.55
	Less-frequent word	16	24.39

Table 1: Average number of iterations needed for all units of homographs and unambiguous words to become saturated in the LH and in the RH networks, when words are presented without context.

from the lexicon. Frequent and less-frequent words (or the dominant versus the subordinate meaning of a given homograph) were selected with a ratio of five to three. Two conditions had to be fulfilled, in order for each network to complete the training process: (a) When presented only with the orthographic features of a given word, the network needed to, successfully, reach a stable state. (b) When presented with the orthographic sub-vector of an ambiguous word together with one group of its semantic features, the network needed to successfully choose the appropriate meaning[5]. After the networks were trained, they were tested by presenting just the orthographic part of the entry as the input (to simulate reading words in isolation). The number of iterations that was needed for all the units of a given word to become saturated (entire vector) was used as an indication of lexical decision times (see Table 3.1 below). Results indicate that when words are presented visually and in isolation, meanings are accessed significantly faster in the LH (see Table 3.1 below). (It is important to note that this pattern always occurred. That is, in each pair of networks, meanings were accessed faster in the LH). In addition, when homographs are presented without context, only the dominant, more frequent meaning is accessed in both networks.

Importantly, Figure 3 shows that although activation of the subordinate meaning is eventually suppressed in both hemispheres, the time course is different in each. Consistent with behavioral data (e.g., Burgess & Simpson [6]), activation of the subordinate meaning in the LH increases more sharply and to a higher degree than in the RH, but then falls more sharply than in the RH. We interpret this as meaning that the secondary possibility remains available for a longer period in the right hemisphere.

This division of labor between the hemispheres (namely, the LH quickly selects one meaning while the RH maintains alternative meanings), also explains

[5] When semantic context was biased toward the subordinate meaning, the LH network was able to satisfy this condition without errors, while the RH made between 2-12% errors .Thus, each network was trained to its optimal level

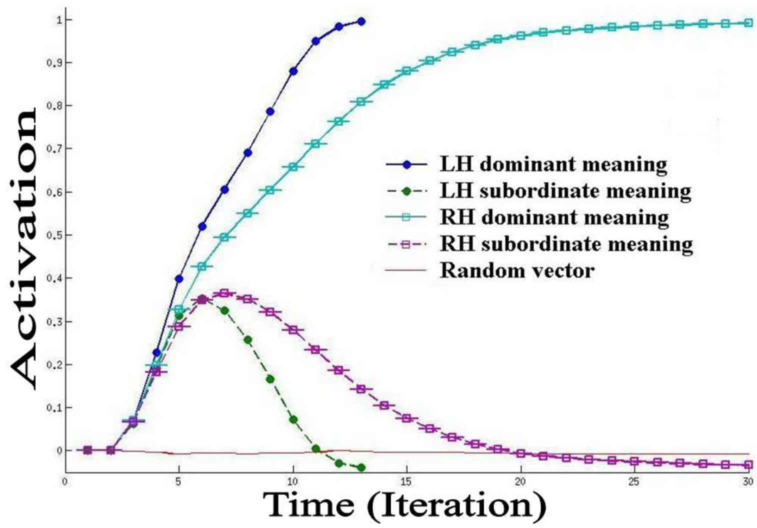

Figure 3: The time course of activation of the subordinate and the dominant meaning of homographs in the LH network and the RH network, when presented in isolation (without context).

why both hemispheres are needed. For example, if it is the case that the LH quickly suppresses the less frequent meaning when biasing contextual information is not available, then it might encounter a problem when a later presented disambiguating context is biased towards the less frequent meaning of the word. In this case, RH contributions may be crucial.

3.2 Step 2: Simulating inter-hemispheric connectivity during lexical ambiguity resolution

We explored the effects of presenting a disambiguating context biased towards the subordinate meaning after the homograph was encountered. Figure 4 (Figure 3 repeated) shows that during a certain time period (see arrow) the right hemisphere, while dynamically on its way to the "attractor" corresponding to the dominant meaning, is less "deep" in the attractor well. We imagine the following scenario. First the networks commence with the orthography. If there is no semantic priming, then they will start the dynamics toward convergence to the dominant attractor. Now assume that during the period of time indicated in Figure 4(see arrow), additional information is given to the network that the other attractor (corresponding to the subordinate meaning) is appropriate. During reading, this might occur when contextual information is presented after encountering the homograph.

In the artificial network, we model this situation by assuming there is new input to the semantic units of the model that biases the results. This was done

Network	Subordinate (appropriate) meaning	Dominant (inappropriate) meaning	Non-convergent
LH network by itself	0%	88%	12%
RH network by itself	63%	17%	20%
LH network + RH information	89%	0%	11%

Table 2: Proportion of subordinate and dominant senses of an ambiguous word accessed, when a subordinately biasing context is given after homograph presentation

by presenting half of the semantic sub-vector consistent with the subordinate meaning when the network had converged to 80% of the dominant solution (about 10 iterations after the orthographic sub-vector was presented). Again, 12 pairs of LH and RH networks were used to simulate 12 subjects in an experiment. We examined how the networks behave under different conditions. First, as a baseline, we examined the individual performance of each network. Then we compared this with the reaction of the LH, if we assume it receives information from the RH. We modeled this situation by simply replacing the values in the LH vector by the values from the RH vector. Results indicate that the most efficient mechanism for "recovery" from erroneous dominant disambiguation is when information is transferred from the RH to the LH (see Table 2 below). Specifically, these simulations show that running the LH without information from the RH results in substantially worse performance. The LH by itself (see Table 2 and Figure 5 below) is unable to recover and erroneously selects the dominant inappropriate meaning, or does not converge at all. The RH model by itself (see Table 2 and Figure 6 below) is more successful than the LH model by itself. Figure 6 shows that in the RH network, both dominant and subordinate meanings were activated. However, performance is perfect when information from the RH is copied into the LH model (see Table 2 and Figure 7 below). Figure 7 shows that under these conditions, the dominant contextually inappropriate meaning that is initially accessed (iteration 10) decreases in activation, while the contextually appropriate subordinate meaning increases in activation until it becomes fully activated (iteration 35).

We see a computational advantage of having these two different networks, in the example where a network has to change after substantial convergence. The results presented here suggest that the LH can converge more quickly than the RH but at the price of loss of information when it has to "change its mind". Fortunately, the different time course in the RH allows the LH to recover by copying its information into its network and then proceeding under the LH.

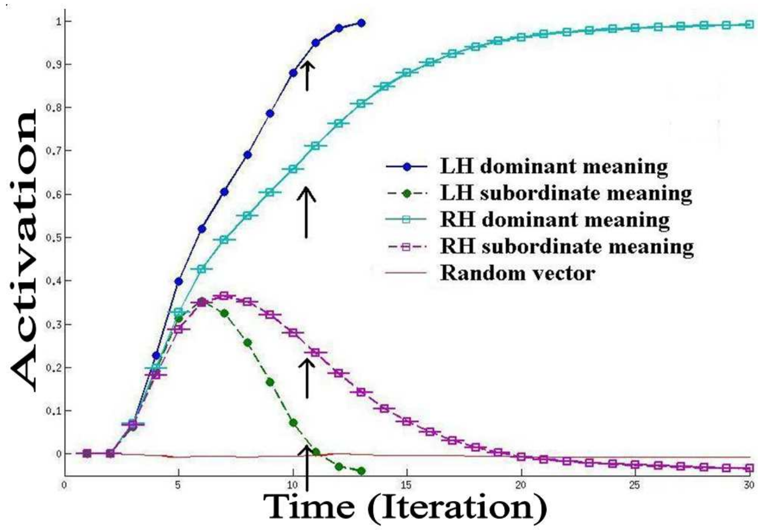

Figure 4: The time course of activation of the subordinate and the dominant meaning of homographs in the LH network and the RH network, when no context is presented. Note that after 11 iterations the subordinate meaning is no longer activated in the LH (see arrow).

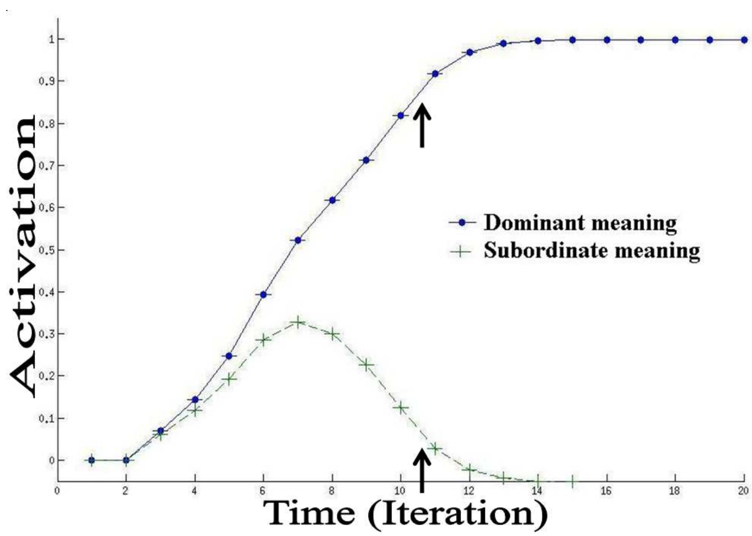

Figure 5: The time course of activation of the subordinate and the dominant meaning of homographs in the LH network, when a subordinately biasing context is given after homograph presentation (see arrow).

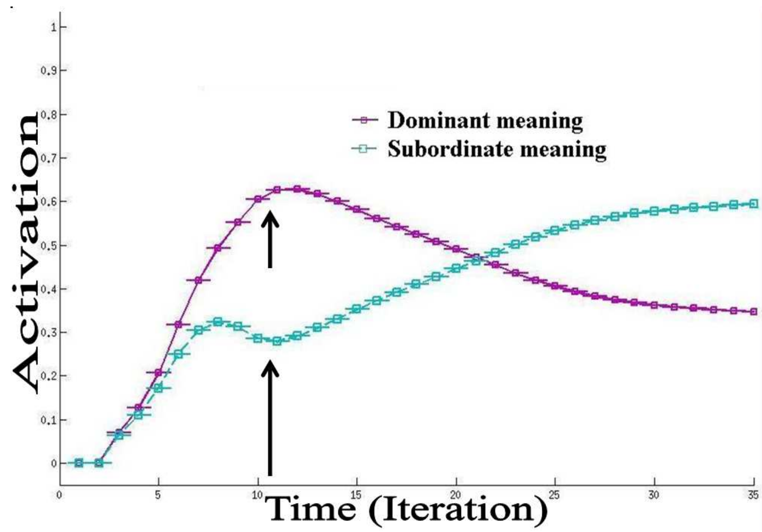

Figure 6: The time course of activation of the subordinate and the dominant meaning of homographs in the RH network when a subordinately biasing context is given after homograph presentation (see arrow).

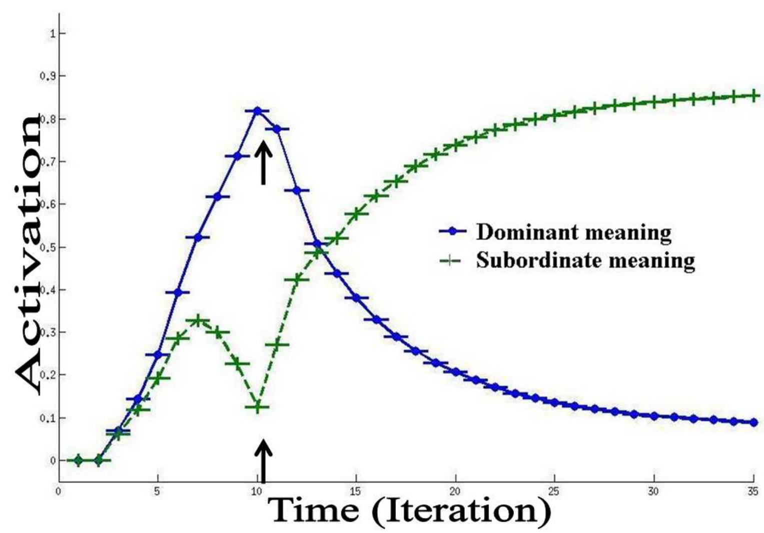

Figure 7: The time course of activation of the subordinate and the dominant meaning of homographs in the LH network, when a subordinately biasing context is given after homograph presentation (see arrow) and information from the RH is copied into the LH model.

3.3 Step 3: Contrasting our model with previous proposals - The disambiguation of homophonic versus heterophonic homographs

Previous proposals are based on evidence from cognitive studies examining the resolution of homophonic homographs (e.g., *bank*). The unvoweled Hebrew, however, offers an opportunity to examine other types of homographs as well. In Hebrew letters represent mostly consonants, and vowels can optionally be superimposed on consonants as diacritical marks. Since the vowel marks are usually omitted, Hebrew readers frequently encounter not only homophonic homographs, but also heterophonic homographs. Both types of homographs have one orthographic representation associated with multiple meanings. They are different however in terms of the relationship between orthography and phonology. In the case of homophonic homographs (*bank*), orthography and phonology are unambiguously related. The phonological route is simple and fast, and may facilitate comprehension. Alternatively, in the case of heterophonic homographs (*bow*), orthography and phonology are ambiguously related. The phonological route is therefore more complicated, and may obstruct comprehension.

In order to contrast the previous proposals (see 1.2, above) with our model we examined the disambiguation of homophonic versus heterophonic homographs in the two hemispheres. That is, if the LH advantage in processing homophonic homographs is due to the LH's unique ability to suppress irrelevant meanings and/or to use contextual information, then a similar advantage should be observed with heterophonic homographs. If, however, the LH advantage in processing homophonic homographs is due to the availability of the phonological route (i.e., direct connection between orthography and phonology), then this advantage may be lost in the case of heterophonic homographs.

Specifically, since the method of training in our model causes frequency to be reflected in the strength of the connections between the units, and because in the LH network we have direct connection between orthography and phonology, we predicted that heterophonic words will have a stronger bias towards the dominant attractor (towards both the dominant phonology and the dominant meaning). On the other hand, in the RH network, there are no direct connections between orthographic and phonological nodes, and we predicted that the training will results in a weaker bias towards the dominant attractor (the contribution of phonology is semantically mediated, and so has less effect). Thus, contrary to the standard explanations, this single architectural difference predicts that when heterophonic homographs are embedded in a subordinately biasing context, it will be the RH which will converge faster towards the appropriate (subordinate) meaning.

To test this prediction, 12 pairs of LH and RH networks were again used to simulate 12 subjects in an experiment. The same lexicon used before was learned by each network: 16 unambiguous words and 16 pairs of homographs (8 homophonic pairs and 8 heterophonic pairs). In order to simulate a biasing context, the networks were tested by presenting the orthographic sub-vector of each homograph, together with part (16 bits) of the semantic sub-vector

		LH	RH
Homophonic Homographs (bank)	Dominant Meaning	8.3	12.1
	Subordinate Meaning	**15.7**	**20.6**
Heterophonic Homographs (bow)	Dominant Meaning	9.5	10.3
	Subordinate Meaning	**23.9**	**18.3**

Table 3: Average number of iterations needed for all units of homophonic and heterophonic homographs to become saturated in the LH and in the RH networks, when part of the semantic sub-vector consistent with the dominant or the subordinate Meaning is presented together with the orthographic sub-vector.

representing its subordinate or dominant meaning.

Results indicate that when homographs were presented with a dominantly-biased context, the LH network was faster (see table 3). However, consistent with our prediction, when a subordinately-biased context was given, the time-line of meaning selection was different for the two types of homographs in the two networks (see table 3 below). For homophones, selection processes were significantly faster in the LH network (see Figure 9 below). In contrast, for heterophones, selection processes were significantly faster in the RH (see Figure 8 below). It is important to note that this "flip" always occurred. That is, in each pair of networks, in the case of homophones (*bank*), the appropriate subordinate meaning was accessed faster in the LH, while in the case of heterophones (*bow*) the appropriate subordinate meaning was accessed faster in the RH).

3.4 A Complementary Human Study

In addition, a behavioral study was conducted in Hebrew and combined a divided visual field (DVF) [6] technique with a semantic priming paradigm. The experimental materials consisted of 112 polarized homographs (both homophonic and heterophonic). Contextual effects were examined by using three different sentential contexts: an ambiguous context ("He went to the *bank*"); dominant-biased context ("The businessman entered the *bank*"); and subordinate-biased

[6] This technique takes advantage of the fact that stimuli presented in the left side of the visual field are initially processed exclusively by the right hemisphere and vice versa. Although information presented in this manner can be later transmitted to both hemisphere, the interpretation of DVF paradigms rests on the assumption that responses to stimuli presented briefly to one visual field reflect mainly the processing of that stimulus by the contralateral hemisphere, so that responses to targets in the right visual field (RVF) reflect left hemisphere (LH) processes and responses to targets in the left visual field (LVF) reflect processes in the right hemisphere (RH). (For theoretical and electrophysiological support for this assumption, see Berardi & Fiorentini [15]; Banich, [1]; Coulson, Federmeier, Van Petten, & Kutas, [7]).

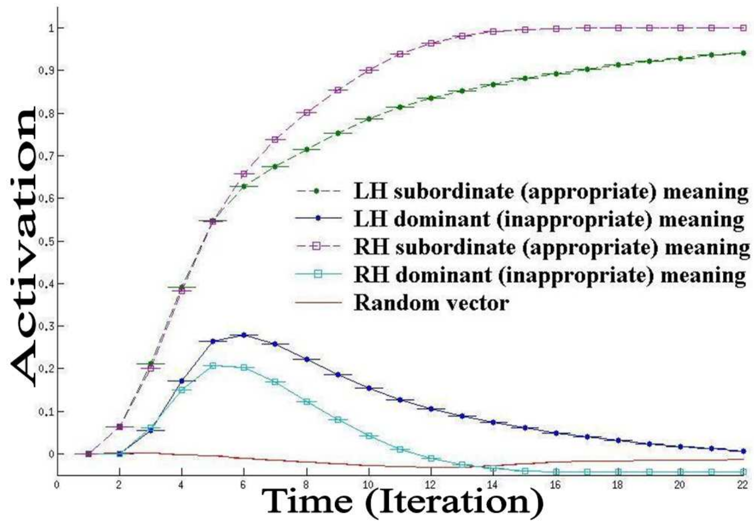

Figure 8: For heterophones, dominant meanings are more difficult to suppress in the LH than in the RH

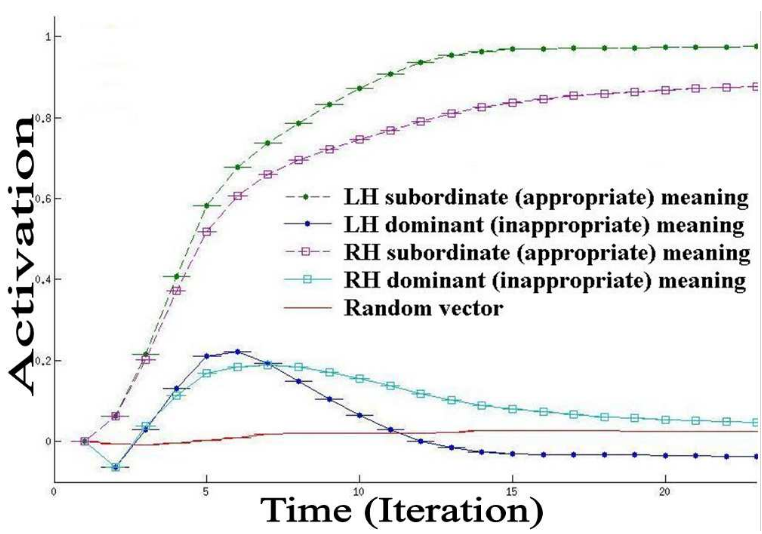

Figure 9: For homophones, dominant meanings are more difficult to suppress in the RH than in the LH

context ("The fisherman sat on the *bank*"). In order to assess the time-line of ambiguity resolution we used two different stimulus onset asynchronies (SOA"s): 250 or 1000 ms.

Subjects were asked to focus on the center of the screen and to silently read sentences that were presented centrally in two stages. First, the sentential context was presented for 1500 ms and then the final ambiguous prime was presented for 150 ms. After the prime disappeared from the screen a target word was presented to the left visual field (LVF) or the right visual field (RVF) for the subject to make a lexical decision. Targets were either related to the dominant or the subordinate meaning or unrelated. Magnitude of priming was calculated by subtracting reaction time (RT) for related targets from RT to unrelated targets. The most interesting results were observed in the subordinate-biasing context condition ('The fisherman sat on the *bank*"): At 250 SOA both meanings (*money* and *river*) were still activated in both hemispheres (Peleg & Eviatar [33]). However, 750 ms later (1000 SOA), we see a different pattern of results in the two visual fields and for the two types of homographs. For homophones (e.g., *bank*), we replicated previous results: The LH selected the contextually appropriate meaning, whereas both meanings were still activated in the RH. In contrast, for heterophones (e.g., *bow*), we get an opposite pattern of results: The LH is unable to suppress the dominant contextually inappropriate meaning, while the RH is able to do so (Peleg and Eviatar [32]). Importantly, this observation fits the results of the computational simulation.

4 General Discussion

Behavioral studies (e.g., Burgess & Simpson [6]; Faust & Gernsbacher, [13]; Faust & Chiarello, [12]) have shown that the time course of lexical ambiguity resolution is different for the left than for the right hemisphere. According to these studies, the LH quickly selects one meaning (e.g., the contextually appropriate meaning), whereas the RH maintains alternative meanings. To account for the sustained activation of contextually incompatible meanings in the RH as opposed to their fast decay in the LH, cognitive researchers (e.g., Beeman [4], [3]; Faust [11]), have suggested that the RH (1) activates a wider range of meanings (2) does not posses a selection mechanism, (3) is less sensitive to sentence-level information, and (4) has a slower activation onset.

In this paper an alternative explanation is presented for the observed hemisphere asymmetries during reading in general and lexical ambiguity resolution in particular. This explanation relates to the different ways in which meanings are accessed in the two hemispheres: While both hemispheres are able to access meaning directly from print, it is only the LH that can directly associate the orthographic representation of a given word with its phonological representations (Zaidel [52], [53], [54]; Iacoboni & Zaidel, [22]). Specifically, the Dual Hemispheric Reading Model was presented in which orthographic, phonological and semantic "neurons" are fully interconnected in the LH (similar to Kawamoto [23]) while in the RH orthographic and phonological neurons are not directly connected.

We tested the model by examining how each network processes two types of polarized homographs: homophonic homographs (e.g., *bank*) and heterophonic homographs (e.g., *bow*). The homographs were either presented in isolation or with context biased toward one interpretation. In all simulations, the dependent variable of interest to us was the time course of response.

In the simulations reported above, it is seen that a single architectural difference between the two networks produces hemispheric asymmetries in the time-course of lexical selection. First, consistent with empirical data, we show that the LH architecture results in faster and more efficient convergence towards the dominant meaning of homographs when the homograph was presented in isolation. Thus, while the LH quickly selects the more frequent alternative, the RH still maintains the subordinate less frequent meaning.

In the second simulation, we explored the effects of presenting a disambiguating context biased towards the subordinate meaning after the homograph was encountered. In this case, the LH was unable to recover because of its fast convergence to the dominant meaning. The RH, on the other hand, because of its slower convergence, is more successful in activating the appropriate subordinate meaning. Significantly, however, the results were optimal when information was transferred from the RH to the LH, and processed within the LH architecture.

This converges with clinical neuropsychological findings that testify to the involvement for both hemispheres in ambiguity resolution (Grindrod & Baum [18]). The LH tendency to select the salient, dominant meaning of an ambiguous word makes it fast, and in most cases, accurate. However, it is less efficient than the RH when a subordinate, less salient interpretation is required. Alternatively, the RH tendency to activate less salient, subordinate meanings alongside the dominant meanings makes it less efficient than the LH in selecting a single alternative, but extremely efficient in situations that require consideration of the less salient meaning. Although other models of hemispheric interaction are possible (e.g., Weems & Reggia [50]), our simulations demonstrate the basic idea that ambiguity resolution requires the intact functioning of *both* hemispheres.

Importantly, this model not only reproduces known human data, it also goes beyond the predictions of previous models proposed by cognitive scientists. As mentioned earlier, empirical studies addressing hemispheric involvement in ambiguity resolution have led to the conclusion that the main difference between the two hemispheres is in their ability to quickly select a single alternative when encountering an ambiguous word. This "standard model" maximizes the LH ability. According to this model, the LH can use both lexical and contextual information, and therefore, in the absence of contextual bias, it quickly selects the salient, more frequent meaning (e.g., Burgess & Simpson [6]), while in the presence of a biased prior context, it quickly selects the contextually appropriate meaning (e.g., Faust & Gernsbacher [13]; Chiarello & Faust [12]). The RH abilities, however, are minimized in this "standard model". It is viewed as less sensitive to meaning pre-dominance or contextual information and therefore maintains alternate meanings regardless of their frequency or contextual appropriateness (e.g., Burgess & Simpson [6]; Faust & Gernsbacher [13]; Chiarello &

Faust [12]).

The results of our third simulation indicate that this "standard model" suggests an asymmetry that is much too strong, if not inaccurate. Instead, it may be posited that both hemispheres can use both frequency and semantic context during ambiguity resolution, but their different architecture leads to a different time course of lexical selection. Importantly, these discrepancies depend on where in the relationship between orthography, phonology and semantics the ambiguity lies. When orthography and phonology is unambiguously related (as in homophonic homographs, e.g., *bank*), lexical selection in a fully connected model (as in the LH network) is faster. However, when orthography and phonology are ambiguously related (as in heterophonic homographs, e.g., *bow*), then a fully connected model may be less efficient. Specifically, because in our model, the method of training causes frequency to be reflected in the strength of the connections between the units, and because in the LH network we have direct connection between orthography and phonology, it turns out that these words have a stronger bias towards the dominant meaning (towards both the dominant phonology and the dominant meaning). On the other hand, in the RH network, there are no direct connections between orthographic and phonological nodes, and so the training results in a weaker bias towards the dominant meaning (the contribution of phonology is semantically mediated, and so has less effect). Thus, in our implementation, the phonology in the RH serves as a "brake" on the convergence of the semantics for both kinds of homographs, however for heterophones there is also the counter-vailing affect of a feedback loop between the semantics and the phonology which overcomes this.

Thus contrary to the standard explanations, this single architectural difference predicts that when heterophonic homographs are embedded in a subordinately biasing context, it will be the RH which will converge faster towards the appropriate (subordinate) meaning.

When investigated, both the simulations and the human behavioral data we gathered show this pattern: dominant inappropriate meanings were activated for a longer period of time in the LH. These results cannot be explained by any of the existing models. Thus, our model not only simulates existing data, but also makes novel predictions which were borne out by our subsequent behavioral studies.

In addition, our model has implications for general theories of visual word recognition, and specifically, for the role that phonology plays in accessing the meaning of words in silent reading. One class of models suggests that printed words activate orthographic codes that are directly related to meanings in semantic memory. An alternative class of models asserts that access to meaning is mediated by phonology (for reviews see Frost [16]; Van Orden and Kloos [49]). In our LH network, orthographic units are directly related to both phonological and semantic units. As a result, meaning activation in the LH is directly influenced by both phonology and orthography. In the RH network, phonological codes are not directly related to orthographic codes and are activated indirectly via semantic codes. This organization results a different sequential ordering of events in which the phonological computation of orthographic representations begins later than the semantic computation of these same representations. As a result, lexical access in the RH is initially more influenced by orthography. This converges with behavioral studies showing that the LH is more influenced by the phonological aspects of a written word (e.g., Zaidel [51]; Zaidel & Peters [55]; Lavidor and Ellis [24]), whereas lexical processing in the RH is more sensitive to the visual form of a written word (e.g., Marsollek, Kosslyn & Squire, [26]; Marsolek, Schacter & Nicholas [27]; Lavidor and Ellis [24]).

The overall picture that emerges from the present results is that hemispheric processes may be more similar than assumed earlier. It seems that both hemispheres have access to the same sources of information (orthographic, phonological, lexical and contextual constraints); however, as a result of the two network architectures, these may be used differently, and with different temporal stages. The idea that RH processing reflects a different pattern of interaction between orthographic phonological and semantic information rather than inability to suppress irrelevant meanings, or to use contextual information, converges with many empirical studies showing RH involvement in comprehending the full meaning of words, phrases and text (e.g., McDonald [30], [31]; Giora, Zaidel, Soroker, Batori, & Kasher, [17]; Federmeier & Kutas, [14]; Coulson & Williams [8]; Mashal, Faust, & Hendler, [28]; Eviatar & Just, [10] 2006).

Taken together, the results of the present study suggest a more coherent picture of how both hemispheres make their unique and critical contribution to language comprehension. Further research is needed to fully explore how the two hemispheres interact during reading comprehension in general and during the resolution of different types of ambiguities in particular.

5 Acknowledgements

This work was supported in part by grant number 956/06 from the Israel Science Foundation (ISF), by the Caesarea Rothschild Institute, the Neurocomputation Laboratory and the Institute of Information Processing and Decision Making of the University of Haifa. The computational results were part of the M.Sc. thesis of Hananel Hazan at the University of Haifa. Portions of this work were presented at the 48th annual meeting of the Psychonomic Society (Los Angeles, 2007) and the WAPCV07 Workshop (Hyderabad, 2007) [34].

References

[1] M.T. Banich. The divided visual field technique in laterality and interhemispheric integration. In *Experimental Methods in Neuropsychology*, pages 47–64. Hughdahl, K (Ed.),New York: Kluwer Publishers., 2003.

[2] M. Beeman, R. Friedman, J. Grafman, E. Perez, S. Diamond, and M. Lindsay. Summation priming and coarse coding in the right hemisphere. *Journal of Cognitive Neuroscience*, 6:26–45, 1994.

[3] M. J. Beeman. Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, 9:512–518, 2005.

[4] M. J. Beeman and C. Chiarello. Approximation by superpositions of a sigmoidal function. *Complementary right- and left-hemisphere language comprehension*, 7(1):2–8, 1998.

[5] C. Burgess and K. Lund. Modeling cerebral asymmetries of semantic memory using high-dimensional semantic space. In *Right Hemisphere Language comprehension: Perspectives from cognitive neuroscience. Hillsdale*. In Beeman M., and Chiarello C. (Eds.), Hillsdale, N.J.: Erlbaum Press, 1998.

[6] C. Burgess and G. B. Simpson. Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language*, 33:86–103, 1988.

[7] S. Coulson, K Federmeier, C. Van Petten, and M. Kutas. Right hemisphere sensitivity to word and sentence-level context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31:129–147, 2005.

[8] S. Coulson and R. W. Williams. Hemispheric asymmetries and joke comprehension. *Neuropsychologia*, 43:128–141, 2005.

[9] S.A. Duffy, R.K. Morris, and K. Rayner. Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, 27:429–446, 1988.

[10] Z. Eviatar and M.A. Just. Brain correlates of discourse processing: An fmri investigation of irony and metaphor comprehension. *Neuropsychologia*, 44:2348–2359, 2006.

[11] M. Faust. Obtaining evidence of language comprehension from sentence priming. In *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience*, pages 161–186. In M.Beeman and C. Chiarello (Eds.),Hillsdale, NJ: Erlbaum., 1998.

[12] M. Faust and C. Chiarello. Sentence context and lexical ambiguity resolution by the two hemispheres. *Neuropsychologia*, 36:827–835, 1998.

[13] M. E. Faust and M. A. Gernsbacher. Cerebral mechanisms for suppression of inappropriate information during sentence comprehension. *Brain and Language*, 53:234–259, 1996.

[14] M. E. Faust and M. A. Gernsbacher. Right words and left words: electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, 8:373–392, 1999.

[15] A. Fiorentini and N. Berardi. Visual perceptual learning: a sign of neural plasticity at early stages of visual processing. *Archives of Italian Biology*, 135:157167, 1997.

[16] R. Frost. Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, 123:71–99, 1998.

[17] R. Giora, E. Zaidel, N. Soroker, G. Batori, and A. Kasher. Differential effect of right and left hemispheric damage on understanding sarcasm and metaphor. *Metaphor and Symbol*, 15:63–83, 2000.

[18] C. M. Grindrod and Baum Shari R. Sensitivity to local sentence context information in lexical ambiguity resolution: Evidence from left- and right-hemisphere-damaged individuals. *Brain and Language*, 85:503–523, 2003.

[19] M. W. Harm and M. S. Seidenberg. Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, 111:662–720, 2004.

[20] G. Hinton and T. Shallice. Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, 98(1):74–95, 1991.

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

[22] M. Iacoboni and E. Zaidel. Hemispheric independence in word recognition: Evidence from unilateral and bilateral presentations. *Brain and Language*, 53:121–140, 1996.

[23] A. H. Kawamoto. Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language*, 32:474–516, 1993.

[24] M . Lavidor and A. W. Ellis. Orthographic and phonological priming in the two cerebral hemispheres. *Laterality*, 8:201–223, 2003.

[25] Larry M. Manevitz and Yigal Zemach. Assigning meaning to data: Using sparse distributed memory for multilevel cognitive tasks. *Neurocomputing*, 14(1):15–39, 1997.

[26] C. J. Marsolek, S. M. Kosslyn, and L. R. Squire. Form-specific visual priming in the right cerebral hemisphere. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18:492–508, 1992.

[27] C. J. Marsolek, D. L. Schacter, and C. D Nicholas. Form-specific visual priming for new associations in the right cerebral hemisphere. *Memory and Cognition*, 24:539–556, 1996.

[28] N. Mashal, M. Faust, and T. Hendler. The role of the right hemisphere in processing nonsalient metaphorical meanings: Application of principal components analysis to fmri data. *Neuropsychologia*, 43(14):2084–2100, 2005.

[29] R.A. Mason and M.A. Just. Lexical ambiguity in sentence comprehension. *Brain Research*, 1146:115–127, 2007.

[30] S. McDonald. Clinical insights into pragmatic theory: Frontal lobe deficits and sarcasm. *Brain and Language*, 53:81–104, 1996.

[31] S. McDonald. Exploring the process of inference generation in sarcasm: A review of normal and clinical studies. *Brain and Language*, 68:486–506, 1999.

[32] O. Peleg and Z. Eviatar. The disambiguation of homophonic versus heterophonic homographs in the two cerebral hemispheres. *A manuscript in preparation.*

[33] O. Peleg and Z. Eviatar. Hemispheric sensitivities to lexical and contextual information: Evidence from lexical ambiguity resolution. *Brain and Language*, page in press, 2008.

[34] O. Peleg, Z. Eviatar, H. Hazan, and L. Manevitz. Differences and interactions between cerebral hemisphers when processing ambiguous homographs. In L. Paletta and E. Rome, editors, *Attention in Cognitive Systems. Theories and Systems from an Interdisciplinary Viewpoint*, volume 4840 of *Lecture Notes in Artificial Intelligence, a subseries of Lecture Notes in Computer Science*. Springer-Verlag, 2008.

[35] O. Peleg, R. Giora, and O. Fein. Salience and context effects: Two are better than one. *Metaphor and Symbol*, 16:173–192, 2001.

[36] O. Peleg, R. Giora, and O. Fein. Contextual strength: The whens and hows of context effects. In *Experimental Pragmatics. Basingstoke: Pagrave.*, pages 172–186. In I. Noveck and D. Sperber (Eds.), 2004.

[37] D. C. Plaut. Lesioning attractor networks as models of neuropsychological deficits. In M. A. Arbib, editor, *The handbook of brain theory and neural networks*, pages 540–543. MIT Press, 1995.

[38] D. C. Plaut. Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language*, 52:25–82, 1996.

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

[40] D. C. Plaut and T. Shallice. Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, 10:377–500, 1993.

[41] M. S. Seidenberg and J. I. McClelland. A distributed, developmental model of word recognition and naming. *Psychological Review*, 96:523–568, 1989.

[42] T.J. Sejnowski and C. R. Rosenberg. Nettalk: a parallel network that learns to read aloud. *Cognitive Science*, 14:179–211, 1986.

[43] G. B. Simpson. Context and the processing of ambiguous words. In *Handbook of psycholinguistics*, pages 359–374. In M. A. Gernsbacher (Ed.), San Diego: Academic Press., 1984.

[44] G. B. Simpson. Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin*, 96:316–340, 1984.

[45] S. I. Small, G. W. Cottrell, and M. K. Tanenhaus. Lexical ambiguity resolution: Perspectives from psycholinguistics. In *neuropsychology, and artificial intelligence*. San Mateo, CA: Morgan Kaufmann, 1988.

[46] J.P. Thivierge, D. Titone, and T.R. Shultz. *Simulating frontotemporal pathways involved in lexical ambiguity resolution.* 2005.

[47] D. A. Titone. Hemispheric differences in context sensitivity during lexical ambiguity resolution. *Brain and Language*, 65:361–394, 1998.

[48] C.A. Tompkins and M.T. Lehman. Interpreting intended meanings after right hemisphere brain damage: An analysis of evidence, potential account, and clinical implications. *Topics in Stroke Rehabilitation*, 5:29–47, 1998.

[49] G. C Van Orden and H. Kloos. The question of phonology and reading. In *The science of reading: A handbook. Blackwell Pub*, pages 39–60. In M. S. Snowling, C. Hulme, and M. Seidenberg (Eds.)., 2005.

[50] S.A. Weems and A. Reggia, J. Hemispheric specialization and independence for word recognition: A comparison of three computational models. *Brain and Language*, 89:554–568, 2004.

[51] E Zaidel. Reading in the disconnected right hemisphere: An aphasiological perspective. In *Dyslexia: Neuronal, Cognitive and Linguistic Aspects*, volume 35, pages 67–91. In: Y. Zotterman, Editor, Oxford, Pergamon Press, 1982.

[52] E Zaidel. Language in the disconnected right hemisphere encyclopedia of neuroscienc. In *Encyclopedia of Neuroscienc.*, volume 89, pages 563–564. Cambridge: Birkhauser, 1987.

[53] E Zaidel. Language functions in the two hemispheres following cerebral commissurotomy and hemispherectomy. In *Handbook of Neuropsychology*, volume 4, pages 115–150. Amsterdam: Elsevier, 1990.

[54] E Zaidel. Language in the right hemisphere following callosal disconnection. In *Handbook of Neuropsychology*, volume 89, pages 369–383. In B. Stemmer and H. Whitaker (Eds.), New York: Academic Press, 1998.

[55] E. Zaidel and A. M. Peters. phonological encoding and ideographic reading by the disconnected right hemisphere: Two case studies. *Brain and Language*, 14:205–234, 1981.