Ann Math Artif Intell (2010) 59:125–147
DOI 10.1007/s10472-010-9210-1

Two hemispheres—two networks: a computational model explaining hemispheric asymmetries while reading ambiguous words

Orna Peleg · Larry Manevitz · Hananel Hazan · Zohar Eviatar

Published online: 17 August 2010

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 closely follows the original computational lines due to Kowamoto (J Mem Lang 32:474–516, 1993) 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

O. Peleg · Z. Eviatar
Institute of Information Processing and Decision Making,
University of Haifa, Haifa, Israel

O. Peleg
e-mail: opeleg@research.haifa.ac.il

Z. Eviatar
e-mail: zohare@research.haifa.ac.il

L. Manevitz (✉) · H. Hazan
Department of Computer Science,
University of Haifa, Haifa, Israel
e-mail: manevitz@cs.haifa.ac.il

H. Hazan
e-mail: hhazan01@cs.haifa.ac.il

reproducing known empirical data, this dual hemispheric reading model makes novel and surprising predictions that were found to be consistent with new human data.

Keywords Disambiguation of natural language • Simulation • Neural networks • Corpus collusum • Modeling • Brain hemispheres

Mathematics Subject Classifications (2010) 68T50 • 92b20 • 68t05 • 68-04

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 et al. [10], Titone [53], Peleg et al. [39, 40], Peleg and Eviatar [37]).

Although effects on ambiguity resolution are still debated (for an overview, see Simpson [48, 49], Small et al. [50]), 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 [27]. 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 [25]). 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 et al. [3], Faust and Gernsbacher [14], Peleg and Eviatar [37]). Moreover, these studies show that the process of ambiguity resolution requires the intact functioning of both cerebral hemispheres (e.g., Grindrod et al. [21], Mason and Just [34]).

Importantly, several studies (e.g., Burgess and Simpson [7], Faust and Gernsbacher [14], Faust and Chiarello [13]) 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, 5], 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 [54]). Another explanation is that the RH is less sensitive to sentence-level information (Faust [12]). As a result, sentential information cannot be used for selection.

Finally, Burgess and Lund [6] 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. Our explanation relates to phonological asymmetries in visual word recognition.

1.2.1 Asymmetries in visual word recognition

Although visual word recognition is normally conceptualized as being driven primarily by the analysis of orthography, it is now commonly accepted that the processing of a printed word is also influenced by information concerning its pronunciation. For example, behavioral studies using the masked-priming paradigm (e.g., [16, 17]) show that target recognition is speeded by the prior brief presentation of a masked pseudo-homophone prime (e.g., koat—COAT) relative to an orthographic control (poat—COAT). This literature has led a number of researchers (e.g., Faust [12]) to suggest that phonological recoding is a mandatory, automatic phase of print processing.

Research on commissurotomy patients, however, suggests that this automatic phonological process proposed by Faust [12] may be an accurate description of reading processes supported by the LH, but may not be applicable to the RH (e.g., Zaidel and Peters [62], Zaidel [58], Baynes and Eliassen [2]). The basic finding, reported by Zaidel and Peters [62], revealed that while the disconnected RH is able to connect the 'sound image of a word' (i.e., its phonological representation) with a picture (i.e., its semantic representation) and to access the meaning of a word from its written form (i.e., its orthographic representation), It is unable to access the phonological form of a word from its written form. The disconnected LH, of course, can access all the representations of the word from its written form.

Hemispheric differences during reading were also investigated in the normal brain.

Many of the studies used divided visual field (DVF) paradigm. This technique takes advantage of the fact that stimuli presented in the left side of the visual field are initially processed exclusively by the RH and vice versa. Although information presented that way can be later transmitted to both hemispheres, the interpretation of DVF studies rests on the assumption that responses to stimuli presented briefly to one visual field reflect mainly the processing of that stimulus by the contra-lateral hemisphere, so that responses to targets in the right visual field (RVF) reflect LH processes and responses to targets in the LVF reflects RH processes (For theoretical and electrophysiological support for this assumption, see Banich [1], Berardi and Fiorentini [18], Coulson et al. [8]).

Similar to the split-brain results, divided visual field studies with intact participants demonstrated that the LH is more influenced by the phonological aspect of written words, whereas word recognition processes in the RH are more influenced by orthography (Lavidor and Ellis [29], Marsolek et al. [31, 32], Baynes and Eliassen [2]). For example, [22] utilized a backward masking paradigm in conjunction with a divided visual field (DVF) display. In that experiment, targets words (e.g., bowl) were presented and backward masked by nonwords that differed in the degree to which they shared orthographic and phonological information with the target. Three types of nonwords were used: pseudo homophone (e.g., bowl—BOAL), orthographically similar, but phonologically less similar (e.g., bowl—BOOL), or unrelated controls (e.g., bowl—MANT). Stimuli were briefly presented to the LVF or to the RVF. Results indicate that responses to targets presented to the RVF/LH, were facilitated in the phonological, pseudo homophone condition relative to the orthographically similar condition. In contrast, responses to targets presented to the LVF/RH, showed a greater degree of facilitation for the orthographically similar condition relative to the unrelated condition. Overall, these observations are consistent with the view that both hemispheres can recognize words visually via orthographic-semantic connections, but orthographic-phonological connections are available only to the LH.

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).

As described above, the visual route is believed to exist in both hemispheres. The phonological route, however, is available only to the left hemisphere (Zaidel [59–61], Iacoboni and Zaidel [26]). 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. (However, see below in “Step 2” for subordinate heterophone discussion.)

We propose a simple model (see Fig. 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 [46]; and see also, Plaut et al. [44], Harm and Seidenberg [23], Thivierge et al. [52]), in the LH, orthographic, phonological and semantic codes are fully interconnected. Importantly, however, in the RH, orthographic and phonological codes are not directly connected.

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.

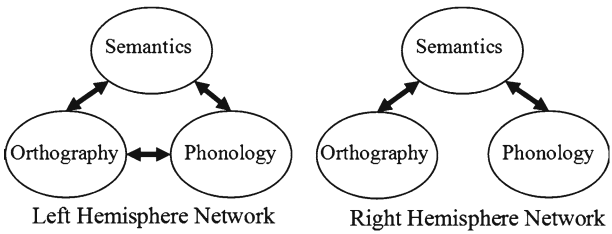


Fig. 1 The dual hemispheric reading model

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 [27] simple recurrent neural network presented above (see Section 1.1). The model includes a LH network and a RH network (see Fig. 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 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); five phonological groups of

[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 [47], Seidenberg and McClelland [46], Plaut [42, 43], Manevitz and Zemach [30], McClelland et al. [44], Plaut and Shallice [45], Hinton and Shallice [24], Kello and Plaut [28].
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 appear in [37, 41].

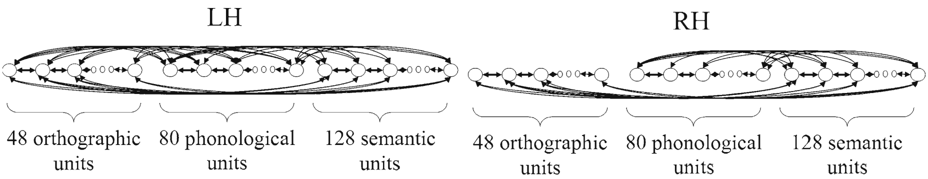


Fig. 2 Architecture of the two connectionist networks. There are 256 units in each network representing the spelling (orthography), pronunciation (phonology) and meaning of words. In the LH network, all units are connected to each other. In the RH network, orthographic and phonological units are not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

features represented its pronunciation (one group for each phoneme)[2] and eight 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.

We proceeded in the following fashion, adapting the method that Kowamoto used in English. Each of orthography, phonology and semantics received its own code. Each code was built up of a concatenation of sub-codes. For orthography, we chose a code for each of the 27 symbols of the Hebrew alphabet. These codes were of 16 bits and carefully chosen so that the hamming distance between any two such codes were the same. Thus for an orthography of three symbols, the code was 48 bits long; and words that were similar orthographically were correspondingly closer in their codes.

For phonology, we assumed a five symbol code as the minimal requirement for Hebrew phonology of a three letter word; resulting in an 80 symbol code for the phonology; where phonologically similar words were correspondingly closer in their codes. The underlying phonological codes were also chosen to be equidistant for each pair under hamming distance.

For semantic encoding, we used an eight symbol code so that all meanings were pairwise hamming equidistant from each other; and these codes had no correlation with either the phonology or the orthography. Thus the total encoding of a word was 256 bits.

The dominant and subordinate meanings were established by the training method, where the dominant/subordinate ratio of training was 10:6. Half of the unambiguous words were trained with the ratio of 10 and the other half with the ratio of 6 as well.

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:

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

[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/ while the Hebrew wordfor "barber" is also spelled "sfr" but pronounced /sapar/.

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 $input_i$ is the net input to unit i. The magnitude of the change in connection strength is determined by the magnitude of the learning constant 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:

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

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] That is, $Limit(x) = x$ if $-1 < x < 1$, has value −1 if $x < -1$, and has value 1 if $x > 1$. (This organization is taken from Kawamoto [27].)

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.[5]

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)

3.1.1 Behavioral data

Behavioral data using the divided visual field technique has shown that when homophonic homographs (e.g., "bank") are presented in a neutral, non-biasing

[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.

[5]50 was chosen as a limit after substantial experimentation did not indicate any significant change in results after allowing much longer time courses.

context, different priming patterns are obtained in the two hemispheres: In the LH, all meanings are immediately activated and shortly afterwards, one meaning is selected on the basis of frequency. The RH, on the other hand, activates all meanings more slowly and maintains these meanings for a longer period of time (e.g., Burgess and Simpson [7], Faust and Gernsbacher [14], Faust and Chiarello [13], Peleg and Eviatar [37, 41]). These patterns are illustrated in Fig. 3 from such behavioral studies in our laboratory. In the RVF/LH (see left panel), both meanings were available at 150 SOA. However, 100 ms. later, only the dominant more frequent meaning remained active. In the LVF/RH (see right panel), the subordinate meaning was activated more slowly, so that 150 ms after the onset of the ambiguous prime, only salient meanings were significantly activated. Shortly afterwards (at 250 SOA), the less-salient meaning was activated alongside the salient one. See [41] for full details.

3.1.2 Simulation

72 pairs of LH and RH networks were used to simulate different 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 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.

The same total number of iterations were used to train both hemispheres. To avoid overtraining, training of a group of resp., homographs, heterographs, and unambiguous was temporarily suspended in any hemisphere when the entire group satisfied the following two conditions: (a) when presented only with the orthographic features of a given word, the network needed to successfully reach the dominant 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

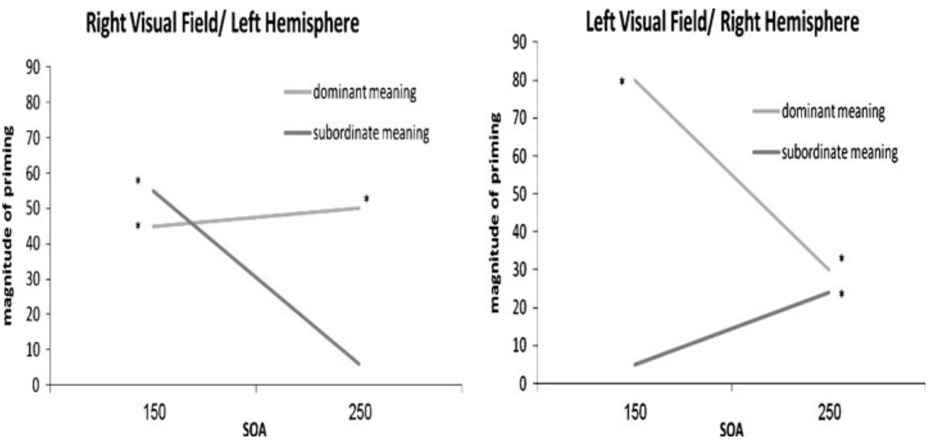


Fig. 3 Magnitude of priming effects (in ms) for targets related to the dominant meaning of homographs (*grey lines*), and to targets related to the subordinate meaning (*black lines*), as a function of SOA (150 or 250 ms), when contexts are kept neutral. Note. *Significant, $p < 0.5$

Table 1 Average number of iterations over 72 pairs of networks 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

		LH network	RH network
Homograph	Dominant/frequent meaning	12.44 (0.65)	15.63 (0.54)
	Subordinate/less-frequent meaning	–	–
Unambiguous word	Frequent word	10.26 (0.30)	11.33 (0.23)
	Less-frequent word	13.30 (1.10)	17.00 (1.40)

The average of the standard deviations is reported. One outlier was omitted from each class

the appropriate meaning. Training was completed on both hemispheres when either hemisphere successfully fulfilled both conditions on all the examples.[6]

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 1). Results indicate that when words are presented visually without context, meanings are accessed significantly faster in the LH. (See Table 1.) This holds for all types of words.

(It is important to note that this pattern always occurred. That is, in each pair of networks, context-free 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.

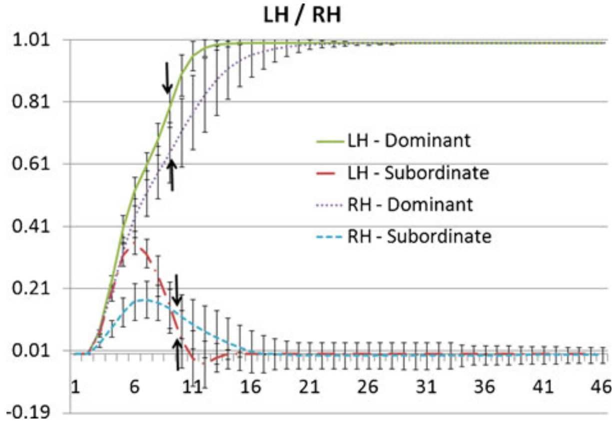
Importantly, Fig. 4 shows that although activation of the subordinate meaning is eventually suppressed in both hemispheres, the time course is different in each. Consistent with behavioral data (e.g., Burgess and Simpson [7]), activation of the subordinate meaning in the LH increases more sharply and to a higher degree than in the RH, but then falls more sharply than in the RH.[7] 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 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. Step 2 addresses this situation.

[6] Under these conditions, the LH always satisfied this condition with no errors; while the RH at this stage still had about 12% errors.

[7] A priori, it could be that this difference is a result only of fewer connections in the RH, and not because of the lack of connectivity between the orthography and phonology. We controlled for this by randomly removing the appropriate number of connections from the LH and rerunning. In this case, while not as fast as before, the LH was still faster than the RH.

Fig. 4 The average 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). Average of standard deviations are indicated on the graph. The *arrows* indicate the time that change of context will be initiated in subsequent graphs (see text)



3.2 Step 2: simulating RH contributions during lexical ambiguity resolution

Recent neuropsychological studies show that ambiguity resolution requires the intact functioning of both hemispheres. For example, not just unilateral LH damage, but also unilateral RH damage leads to deficits in ambiguity resolution (e.g., Grindrod and Baum [21]). Similarly, imaging studies reveal bilateral activation during ambiguity resolution (e.g., Mason and Just [34]). However, the unique contribution of each hemisphere to reading in general and to the resolution of homographs in particular remains to be elucidated. Step 2 explores one possible situation where hemispheric sharing of information results in better processing outcomes, supporting the hypothesis that both hemispheres are needed in successful language processing. We explored the effects of presenting a disambiguating context biased towards the subordinate meaning after the homograph was encountered. Figure 4 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 at the time indicated in Fig. 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 by presenting half of the semantic sub-vector consistent with the subordinate meaning when the LH network had converged to 80%[8] of the dominant solution (about 9–

[8]80% was chosen after much experimentation. It was the latest point in the LH convergence when the recovery phenomenom described below could be observed. Results were comparable when the LH was only 65% convergent. Of course, LH could not converge at all when the RH was at the 80% level.

Table 2 Proportion of subordinate and dominant senses of an ambiguous word accessed, when a subordinately biasing context is given subsequent to the original homograph presentation at the time when the LH has converged 80% of the way to the dominant meaning

Network	Subordinate (appropriate) meaning (%)	Dominant (inappropriate) meaning (%)	Non-convergent (%)
LH network by itself	0.00	87.23	12.77
RH network by itself	53.21	21.35	25.54
LH network + RH information	81.42	0.00	18.58

11 iterations after the orthographic sub-vector was presented). Again, 72 pairs of LH and RH networks were used to simulate 72 subjects in an experiment. Note that the responses of different networks might differ because they are at different depths of the appropriate attractor. 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.[9] 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). 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, line 1 and Fig. 5) was not able to recover and erroneously selects the dominant inappropriate meaning, or does not converge at all. The RH model by itself (see Table 2, line 2 and Fig. 5) is more successful than the LH model by itself. Figure 5 shows that in the RH network, both dominant and subordinate meanings were activated both for longer (up to about 13 iterations) before they commit to the choice of a meaning. From the Table 2, line 2, we see that the RH is able to successfully recover and converge to the subordinate meaning for a substantial percentage of cases. However, performance is greatly improved with no convergence at all to the inappropriate meaning when information from the RH is copied into the LH model (see Table 2, line 3 and Fig. 6). Figure 6 shows that under these conditions, the dominant contextually inappropriate meaning that is initially accessed (where indicated by the arrow) decreases in activation, while the contextually appropriate subordinate meaning increases in activation until it becomes fully activated. (As a control, we also copied the RH into the LH and, as expected, this only sets the LH "back in time" and it continues to converge to the dominant solution.)

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.

[9]It is important to note, that we do not see this as a model of callosal connectivity, but rather, as the beginning of an exploration of the manner in which transfer of information can affect these time courses.

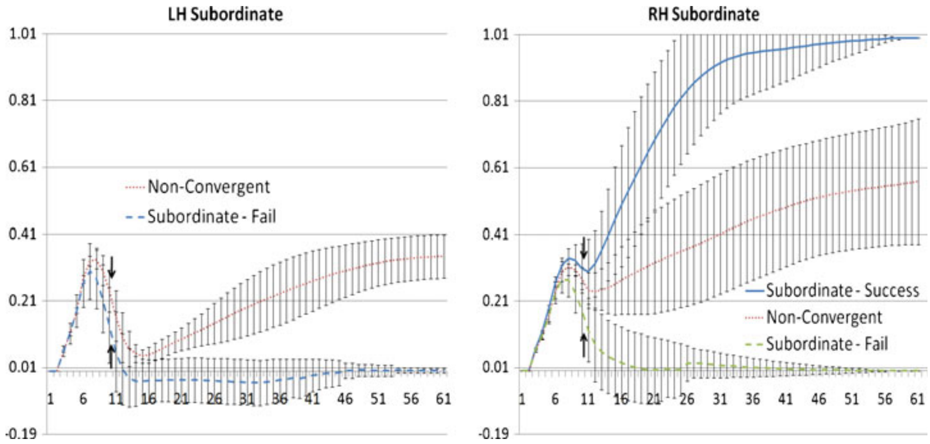
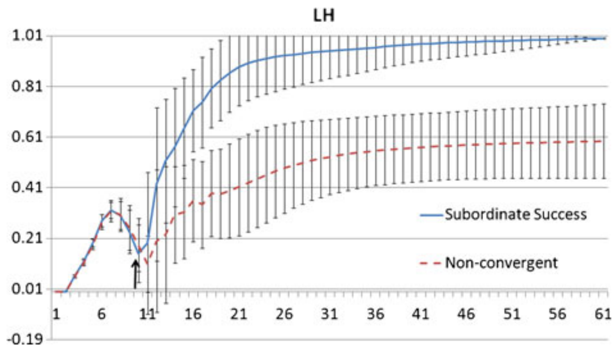


Fig. 5 The time course of activation of the subordinate and the dominant meaning of homographs in the LH and RH networks, when a subordinately biasing context is given after homograph presentation (see *arrow*). "Subordinate-Fail" means converged to the dominant meaning. Note that the LH continues to converge to the dominant meaning

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

Fig. 6 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



are ambiguously related. The phonological route is therefore more complicated, and may obstruct comprehension.

As mentioned earlier, previous studies suggest that orthographic-semantic connections exist in both hemispheres, whereas orthographic-phonological direct associations are available only to the LH (e.g., Zaidel and Peters [62], Lavidor and Ellis [29], Smolka and Eviatar [51]).

On the basis of these findings, the model described in this paper was developed in which both hemispheres exploit orthographic, phonological, and semantic information in the processing of written words. However, in the LH, orthographic, phonological, and semantic representations are fully interconnected, while there are no direct connections between phonological and orthographic units in the RH. We claim that this single difference in hemispheric functional architecture results in hemisphere asymmetries in the disambiguation of homographs in particular, and more broadly, in the processing of written words. According to our proposed model, when orthographic and phonological representations are unambiguously related (as in the case of homophonic homographs like bank), meaning activation is faster in the LH than in the RH, because all related meanings are immediately boosted by both orthographic and phonological sources of information. However, when a single orthographic representation is associated with multiple phonological representations, (as in the case of heterophonic homographs like bow) meanings may be activated more slowly in LH than in the RH, due to the competition between the different phonological alternatives.

In order to contrast the previous proposals (see Section 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, 72 pairs of LH and RH networks were again used to simulate 72 subjects in an experiment. The same lexicon used before was learned by each network. 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 representing its subordinate or dominant meaning.

Table 3 Average number of iterations over 72 pairs of networks 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

		LH	RH
Homophonic homographs (bank)	Dominant meaning	7.28 (0.17)	9.60 (0.04)
	Subordinate meaning	8.64 (0.43)	12.03 (0.60)
Heterophonic homographs (bow)	Dominant meaning	7.65 (0.17)	8.11 (0.18)
	Subordinate meaning	10.18 (0.89)	9.82 (0.38)

Results indicate that when homographs were presented with a dominantly-biased context, the LH network was significantly faster. (See Table 3). However, consistent with our prediction, when a subordinately-biased context was given, the time of meaning selection was different for the two types of homographs in the two networks. (See Table 3). For homophones, selection processes were significantly faster in the LH network. (See Fig. 7 (left side) and Table 3, line 2.) In contrast, for heterophones, there was no significant difference between RH and LH. (See Fig. 7 (right side) and Table 3, line 4.)

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 always accessed faster or the same speed in the RH).

In addition (see Fig. 8, during the selection process, the inappropriate (dominant) meaning stays more available in the RH than the LH during the convergence to the subordinate homophone context. On the other hand and in contrast (see Fig. 8, during the subordinate heterophone context, the inappropriate (dominant) meaning stays more available in the LH. This is in parallel to the human experiments, discussed in the following section.

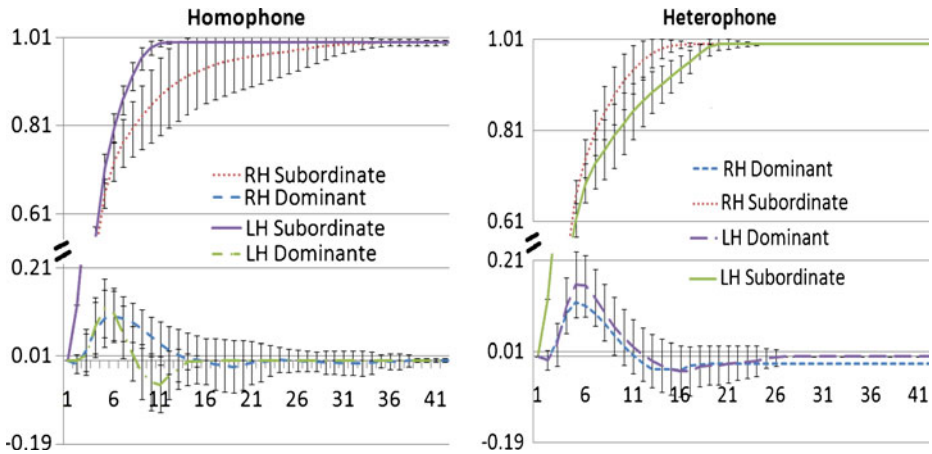


Fig. 7 *Left* For homophones, the LH converges faster to the subordinate meaning. *Right* For heterophones, the RH converges faster to the subordinate meaning. See also Fig. 8 and compare Fig. 9

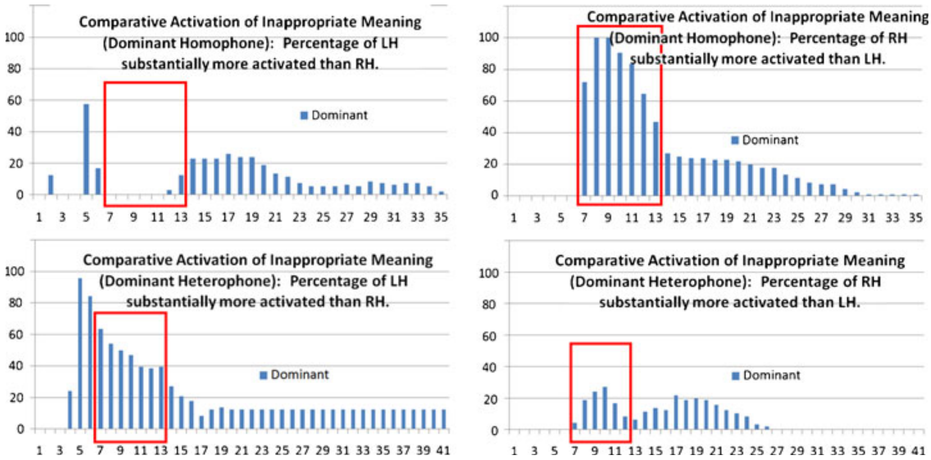


Fig. 8 Comparative activations. *Bars* measure the percentage of words where activation of one hemisphere is more than 0.02 (of full activation) higher than the other hemisphere. Iterations 7–13 (*boxed*) correspond to the time when the inappropriate activations are decaying in both hemispheres. *Top row* For homophones, dominant meanings are more difficult to suppress in the RH than in the LH. *Bottom row* For heterophones, dominant meanings are more difficult to suppress in the LH than in the RH. This coincides with the human results. Compare Fig. 9

3.3.1 A complementary human study

In our studies [37] conducted after the above simulation, a DVF technique was employed in conjunction with the lexical-priming paradigm. Participants were asked to silently read sentences that ended with either homophonic or heterophonic homographs and to perform a lexical decision task on targets presented laterally (to the LVF or to the RVF) 250 or 1,000 ms after the onset of the final homograph. Sentential contexts were biased towards the subordinate interpretation of the final homograph. Targets were either related to one of the meanings of the ambiguous prime, or unrelated. Magnitude of priming was calculated by subtracting reaction time (RT) to related targets from RT to unrelated targets. Translated examples of the stimuli in the different conditions are presented in Table 4 and the pattern of results is shown in Fig. 9.

In addition, a behavioral study was conducted in Hebrew and combined a divided visual field (DVF) 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

Table 4 Translated examples of stimuli

Homograph	Sentence context	Homograph	Pronunciation	Target words
Homophonic	Subordinate:	contract	/XOZE/	Dom: document
	The children of Israel listened to the	seer		Sub: prophet
Heterophonic	Subordinate:	book	/SEFER/	Dom: reading
	The bride made an appointment with the	barber	/SAPAR/	Sub: hair

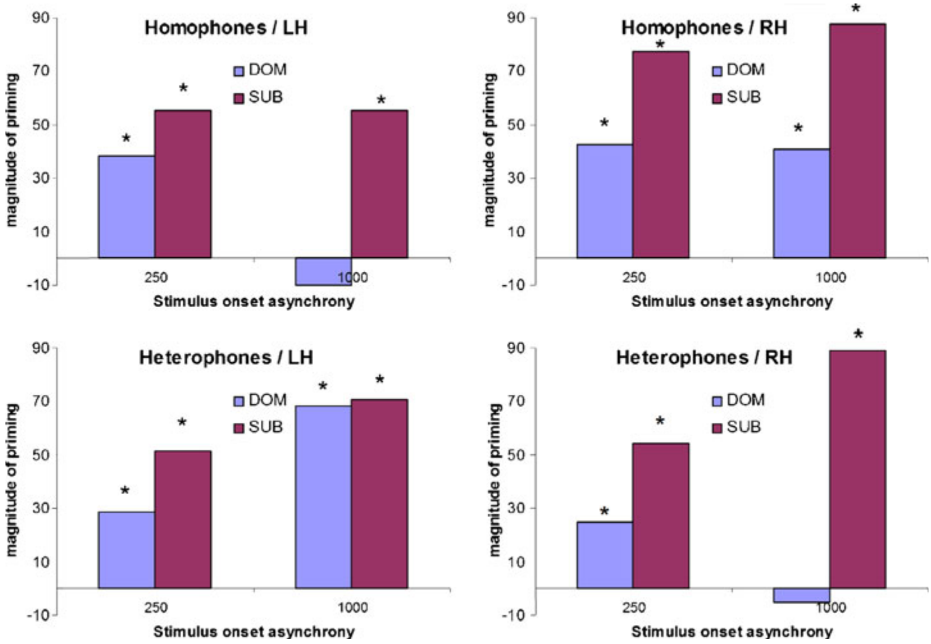


Fig. 9 Behavioral study: for homophones, we replicated previous human results. For heterophones, it is more difficult for the LH to suppress the dominant contextually inappropriate meaning. This corresponds to the computational simulation result in Step 3 and expressed in Table 3

contexts: an ambiguous context ("He went to the *bank*"); dominant-biased context ("The businessman entered the *bank*"); and subordinate-biased 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 1,000 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 1,500 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 and Eviatar [37]). However, 750 ms later (1,000 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 [41]). Importantly, this observation fits the results of the above computational simulation.

4 General discussion

Behavioral studies (e.g., Burgess and Simpson [7], Faust and Gernsbacher [14], Faust and Chiarello [13]) 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, 5], Faust [12]), 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 [59–61], Iacoboni and Zaidel [26]). Specifically, the Dual Hemispheric Reading Model was presented in which orthographic, phonological and semantic "neurons" are fully interconnected in the LH (similar to Kawamoto [27]) 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 and Baum [21]). 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 and Reggia [56]), 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 and Simpson [7]), while in the presence of a biased prior context, it quickly selects the contextually appropriate meaning (e.g., Faust and Gernsbacher [14], Chiarello and Faust [13]). 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 and Simpson [7], Faust and Gernsbacher [14], Chiarello and Faust [13]).

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 [19], Van Orden and Kloos [55]). 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 [57], Zaidel and Peters [62], Lavidor and Ellis [29]), whereas lexical processing in the RH is more sensitive to the visual form of a written word (e.g., Marsolek et al. [31, 32], Lavidor and Ellis [29]).

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 [35, 36], Giora et al. [20], Federmeier and Kutas [15], Coulson and Williams [9], Mashal et al. [33], Eviatar and Just [11]).

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.

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) [38].

References

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

2. Baynes, K., Eliassen, J.C.: The visual lexicon: its access and organization in commissurotomy patients. Right hemisphere language comprehension: perspectives from cognitive neuroscience. In: Beeman, M., Chiarello, C. (eds.) Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience. Erlbaum, Mahwah (1998)
3. Beeman, M., Friedman, R., Grafman, J., Perez, E., Diamond, S., Lindsay, M.: Summation priming and coarse coding in the right hemisphere. J. Cogn. Neurosci. **6**, 26–45 (1994)
4. Beeman, M.J.: Bilateral brain processes for comprehending natural language. Trends Cogn. Sci. **9**, 512–518 (2005)
5. Beeman, M.J., Chiarello, C.: Approximation by superpositions of a sigmoidal function. Complementary right and left hemisphere language comprehension. Curr. Dir. Psychol. Sci. **7**, 2–8 (1998)
6. Burgess, C., Lund, K.: Modeling cerebral asymmetries of semantic memory using high-dimensional semantic space. In: Beeman, M., Chiarello, C. (eds.) Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience. Erlbaum, Hillsdale (1998)
7. Burgess, C., Simpson, G.B.: Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. Brain Lang. **33**, 86–103 (1988)
8. Coulson, S., Federmeier, K., Van Petten, C., Kutas, M.: Right hemisphere sensitivity to word and sentence-level context. J. Exper. Psychol., Learn., Mem., Cogn. **31**, 129–147 (2005)
9. Coulson, S., Williams, R.W.: Hemispheric asymmetries and joke comprehension. Neuropsychologia **43**, 128–141 (2005)
10. Duffy, S.A., Morris, R.K., Rayner, K.: Lexical ambiguity and fixation times in reading. J. Mem. Lang. **27**, 429–446 (1988)
11. Eviatar, Z., Just, M.A.: Brain correlates of discourse processing: an FMRI investigation of irony and metaphor comprehension. Neuropsychologia **44**, 2348–2359 (2006)
12. Faust, M.: Obtaining evidence of language comprehension from sentence priming. In: Beeman, M., Chiarello, C. (eds.) Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience, pp. 161–186. Erlbaum, Hillsdale (1998)
13. Faust, M., Chiarello, C.: Sentence context and lexical ambiguity resolution by the two hemispheres. Neuropsychologia **36**, 827–835 (1998)
14. Faust, M.E., Gernsbacher, M.A.: Cerebral mechanisms for suppression of inappropriate information during sentence comprehension. Brain Lang. **53**, 234–259 (1996)
15. Faust, M.E., Gernsbacher, M.A.: Right words and left words: electrophysiological evidence for hemispheric differences in meaning processing. Cogn. Brain Res. **8**, 373–392 (1999)
16. Ferrand, L., Grainger, J.: Phonology and orthography in visual word recognition: evidence from masked nonword priming. Q. J. Exp. Psychol. **45A**, 353–372 (1992)
17. Ferrand, L., Grainger, J.: The time-course of phonological and orthographic code activation in the early phases of visual word recognition. Bull. Psychon. Soc. **31**, 119–122 (1993)
18. Fiorentini, A., Berardi, N.: Visual perceptual learning: a sign of neural plasticity at early stages of visual processing. Arch. Ital. Bio. **135**, 157–167 (1997)
19. Frost, R.: Toward a strong phonological theory of visual word recognition: true issues and false trails. Psychol. Bull. **123**, 71–99 (1998)
20. Giora, R., Zaidel, E., Soroker, N., Batori, G., Kasher, A.: Differential effect of right and left hemispheric damage on understanding sarcasm and metaphor. Metaphor Symb. **15**, 63–83 (2000)
21. Grindrod, C.M., Baum Shari, R.: Sensitivity to local sentence context information in lexical ambiguity resolution: evidence from left- and right-hemisphere-damaged individuals. Brain Lang. **85**, 503–523 (2003)
22. Halderman, L.K., Chiarello, C.: Cerebral asymmetries in early orthographic and phonological reading processes: evidence from backward masking. Brain Lang. **95**, 342–352 (2005)
23. Harm, M.W., Seidenberg, M.S.: Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. Psychol. Rev. **111**, 662–720 (2004)
24. Hinton, G., Shallice, T.: Lesioning an attractor network: investigations of acquired dyslexia. Psychol. Rev. **98**(1), 74–95 (1991)
25. Hopfield, J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA **79**, 2554–2558 (1982)
26. Iacoboni, M., Zaidel, E.: Hemispheric independence in word recognition: evidence from unilateral and bilateral presentations. Brain Lang. **53**, 121–140 (1996)
27. Kawamoto, A.H.: Nonlinear dynamics in the resolution of lexical ambiguity: a parallel distributed processing account. J. Mem. Lang. **32**, 474–516 (1993)

28. Kello, C.T., Plaut, D.C.: Strategic control over rate of processing in word reading: a computational investigation. J. Mem. Lang. **48**, 207–232 (2003)
29. Lavidor, M., Ellis, A.W.: Orthographic and phonological priming in the two cerebral hemispheres. Laterality **8**, 201–223 (2003)
30. Manevitz, L.M., Zemach, Y.: Assigning meaning to data: using sparse distributed memory for multilevel cognitive tasks. Neurocomputing **14**(1), 15–39 (1997)
31. Marsolek, C.J., Kosslyn, S.M., Squire, L.R.: Form-specific visual priming in the right cerebral hemisphere. J. Exper. Psychol., Learn., Mem., Cogn. **18**, 492–508 (1992)
32. Marsolek, C.J., Schacter, D.L., Nicholas, C.D.: Form-specific visual priming for new associations in the right cerebral hemisphere. Mem. Cogn. **24**, 539–556 (1996)
33. Mashal, N., Faust, M., Hendler, T.: 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)
34. Mason, R.A., Just, M.A.: Lexical ambiguity in sentence comprehension. Brain Res. **1146**, 115–127 (2007)
35. McDonald, S.: Clinical insights into pragmatic theory: frontal lobe deficits and sarcasm. Brain Lang. **53**, 81–104 (1996)
36. McDonald, S.: Exploring the process of inference generation in sarcasm: a review of normal and clinical studies. Brain Lang. **68**, 486–506 (1999)
37. Peleg, O., Eviatar, Z.: Hemispheric sensitivities to lexical and contextual constraints: evidence from ambiguity resolution. Brain Lang. **105**(2), 71–82 (2008)
38. Peleg, O., Eviatar, Z., Hazan, H., Manevitz, L.: Differences and interactions between cerebral hemisphers when processing ambiguous homographs. In: Paletta, L., Rome, E. (eds.) Attention in Cognitive Systems. Theories and Systems from an Interdisciplinary Viewpoint. Lecture Notes in Artificial Intelligence, vol. 4840. A Subseries of Lecture Notes in Computer Science. Springer (2008)
39. Peleg, O., Giora, R., Fein, O.: Salience and context effects: two are better than one. Metaphor Symb. **16**, 173–192 (2001)
40. Peleg, O., Giora, R., Fein, O.: Contextual strength: the whens and hows of context effects. In: Noveck, I., Sperber, D. (eds.) Experimental Pragmatics, pp. 172–186. Pagrave, Basingstoke (2004)
41. Peleg, O., Eviatar, Z.: Semantic asymmetries are modulated by phonological asymmetries: evidence from the disambiguation of homophonic versus heterophonic homographs. Brain Cogn. **70**, 154–162 (2009)
42. Plaut, D.C.: Lesioning attractor networks as models of neuropsychological deficits. In: Arbib, M.A. (ed.) The Handbook of Brain Theory and Neural Networks, pp. 540–543. MIT Press (1995)
43. Plaut, D.C.: Relearning after damage in connectionist networks: toward a theory of rehabilitation. Brain Lang. **52**, 25–82 (1996)
44. Plaut, D.C., McClelland, J.L., Seidenberg, M.S., Patterson, K.: Understanding normal and impaired word reading: computational principles in quasi-regular domains. Psychol. Rev. **103**, 56–115 (1996)
45. Plaut, D.C., Shallice, T.: Deep dyslexia: a case study of connectionist neuropsychology. Cogn. Neuropsychol. **10**, 377–500 (1993)
46. Seidenberg, M.S., McClelland, J.I.: A distributed, developmental model of word recognition and naming. Psychol. Rev. **96**, 523–568 (1989)
47. Sejnowski, T.J., Rosenberg, C.R.: Nettalk: a parallel network that learns to read aloud. Cogn. Sci. **14**, 179–211 (1986)
48. Simpson, G.B.: Context and the processing of ambiguous words. In: Gernsbacher, M.A. (ed.) Handbook of Psycholinguistics, pp. 359–374. Academic, San Diego (1984)
49. Simpson, G.B.: Lexical ambiguity and its role in models of word recognition. Psychol. Bull. **96**, 316–340 (1984)
50. Small, S.I., Cottrell, G.W., Tanenhaus, M.K.: Lexical ambiguity resolution: perspectives from psycholinguistics. In: Neuropsychology, and Artificial Intelligence. Morgan Kaufmann, San Mateo (1988)
51. Smolka, E., Eviatar, Z.: Phonological and orthographic visual word recognition in the two cerebral hemispheres: evidence from Hebrew. Cogn. Neuropsychol. **23**(6), 972–989(18) (2006)
52. Thivierge, J.P., Titone, D., Shultz, T.R.: Simulating Frontotemporal Pathways Involved in Lexical Ambiguity Resolution (2005)
53. Titone, D.A.: Hemispheric differences in context sensitivity during lexical ambiguity resolution. Brain Lang. **65**, 361–394 (1998)

54. Tompkins, C.A., Lehman, M.T.: Interpreting intended meanings after right hemisphere brain damage: an analysis of evidence, potential account, and clinical implications. Top. Stroke Rehabil. **5**, 29–47 (1998)
55. Van Orden, G.C., Kloos, H.: The question of phonology and reading. In: Snowling, M.S., Hulme, C., Seidenberg, M. (eds.) The Science of Reading: A Handbook, pp. 39–60. Blackwell (2005)
56. Weems, S.A., Reggia, J.A.: Hemispheric specialization and independence for word recognition: a comparison of three computational models. Brain Lang. **89**, 554–568 (2004)
57. Zaidel, E.: Reading in the disconnected right hemisphere: an aphasiological perspective. In: Zotterman, Y. (ed.) Dyslexia: Neuronal, Cognitive and Linguistic Aspects, vol. 35, pp. 67–91. Pergamon, Oxford (1982)
58. Zaidel, E.: The Dual Brain. Hemispheric Specialization in Humans. The Guildford Press, New York (1985)
59. Zaidel, E.: Language in the disconnected right hemisphere encyclopedia of neuroscience. In: Encyclopedia of Neuroscience, vol. 89, pp. 563–564. Birkhauser, Cambridge (1987)
60. Zaidel, E.: Language functions in the two hemispheres following cerebral commissurotomy and hemispherectomy. In: Handbook of Neuropsychology, vol. 4, pp. 115–150. Elsevier, Amsterdam (1990)
61. Zaidel, E.: Language in the right hemisphere following callosal disconnection. In: Stemmer, B., Whitaker, H. (eds.) Handbook of Neuropsychology, vol. 89, pp. 369–383. Academic, New York (1998)
62. Zaidel, E., Peters, A.M.: Phonological encoding and ideographic reading by the disconnected right hemisphere: two case studies. Brain Lang. **14**, 205–234 (1981)