Hemispheric asymmetry in perception: A differential encoding account

Janet H. Hsiao[1], Ben Cipollini[2], & Garrison W. Cottrell[3]

[1]Department of Psychology, University of Hong Kong, Hong Kong

[2]Department of Cognitive Science, University of California San Diego, USA

[3]Department of Computer Science & Engineering, University of California San Diego, USA

Address for correspondence:

Janet H. Hsiao

Department of Psychology

University of Hong Kong

Phone: +852 3917 4874

Email: jhsiao@hku.hk

ABSTRACT

Hemispheric asymmetry in the processing of local and global features has been argued to originate from differences in frequency filtering in the two hemispheres, with little neurophysiological support. Here we test the hypothesis that this asymmetry takes place at an encoding stage beyond the sensory level, due to asymmetries in anatomical connections within each hemisphere. We use two simple encoding networks with differential connection structures as models of differential encoding in the two hemispheres, based upon a hypothesized generalization of neuroanatomical evidence from the auditory modality to the visual modality: the connection structure between columns is more distal in the language areas of the left hemisphere, and more local in the homotopic regions in the right hemisphere. We show that both processing differences and differential frequency filtering can arise naturally in this neurocomputational model with neuroanatomically-inspired differences in connection structures within the two model hemispheres, suggesting that hemispheric asymmetry in the processing of local and global features may be due to hemispheric asymmetry in connection structure rather than in frequency tuning.

Hemispheric asymmetry in perception: A differential encoding account

INTRODUCTION

How the brain processes visual stimuli at the global and local level has been extensively examined. Navon (1977) proposed the "global precedence hypothesis" and argued that the global form of a visual stimulus is unavoidably recognized before the local forms. This effect was later shown to depend on both the characteristics of the local and global forms and the hemispheric asymmetry in the perception of local and global features (Hoffman, 1980). Follow-up studies further confirmed that there is a right visual field (RVF)/left hemisphere (LH) advantage for responses to local features and a left visual field (LVF)/right hemisphere (RH) advantage for responses to global features (e.g., Martin, 1979; Sergent, 1982; Ivry & Robertson, 1998; Delis, Robertson, & Efron, 1986; Robertson & Delis, 1986; Van Kleeck, 1989; Robertson, Lamb, & Zaidel, 1993; Martinez et al., 1997; Proverbio, Minniti, & Zani, 1998; Han et al., 2002; Weissman & Woldorff, 2005; Flevaris, Bentin, & Robertson, 2010). For example, by using hierarchical letter stimuli due to Navon (1977), where a large letter is made up of many smaller letters (Figure 1a), it has been shown that subjects are faster at detecting small letters when they are presented to the RVF/LH, and faster at detecting large letters when presented to the LVF/RH (e.g., Sergent, 1982; Ivry & Robertson, 1998; Van Kleeck, 1989; Figure 1b). Accordingly, Sergent (1982) concluded that global precedence in form analysis is a property of the RH but not the LH. She referred to the two levels of visual stimuli as having differential spatial frequency contents: low frequency for global features and high frequency for local features, and argued that the LH is more adept in processing high frequency information, whereas the RH is more efficient in processing low frequency information. This differential frequency processing account was supported by some follow-up studies (Ivry & Robertson, 1998), using tasks such as spatial frequency identification (Kitterle, Christman, & Hellige, 1990) and discrimination (Proverbio,

Zani, & Avella, 2002), face recognition (Keenan, Whitman, & Pepe, 1989), and in fMRI (Han et al., 2002) and EEG (Flevaris, Bentin, & Robertson, 2011) studies.

A fundamental problem with this proposal is that studies examining grating detection do not support hemispheric specialization for particular frequency ranges (e.g., Kitterle, Christman, & Hellige, 1990; Rijsdijk, Kroon, &Van der Wildt, 1980; Di Lollo, 1981; Peterzell, Harvey, & Hardyck, 1989; Fendrich & Gazzaniga, 1990; Peterzell, 1991). For example, Peterzell, Harvey, and Hardyck (1989) presented vertical gratings to the LVF and the RVF of the participants and found no difference between the two hemispheres in contrast-sensitivity functions measured or in visible persistence durations. Fendrich and Gazzaniga (1990) presented a pair of Gaussian windowed sinusoidal gratings either within the LVF and the RVF of both commissurotomy patients and healthy controls, and asked them to judge whether the pair had the same orientation; they showed that there was no indication of an interaction between visual field and spatial frequency of the gratings in this task (Ivry and Robertson (1998) have argued that this result is due to the use of absolute rather than relative frequencies in these studies). Sergent (1982) thus argued that this asymmetry "must result from processing taking place beyond the sensory level". Consistent with this speculation, an ERP study found that the hemispheric asymmetry in processing global vs. local information was observed in the N2 component but not in the earlier, sensory-evoked P1 component, suggesting a higher stage of perceptual processing (Heinze, Hinrichs, Scholz, Burchert, & Mangun, 1998). fMRI studies have also shown that activation corresponding to the observed behavioral asymmetry was found in the occipitotemporal regions of the two hemispheres (Martinez et al., 1997).

A similar hemispheric asymmetry has also been consistently reported in auditory perception. For example, in dichotic listening studies of speech recognition, it has been shown that there is an advantage for responses to prosody, which relies more on low frequency information, when the stimulus is presented to the left ear/RH, and an advantage for responses to content, which relies more on high frequency information, when the stimulus is presented to the right ear/LH (e.g., Ivry & Lebby,

1993; Ivry & Robertson, 1998; Ley & Bryden, 1982; Bartholomeus, 1974; Chan & Hsiao, in press). In addition, similar to visual processing, it has been proposed that auditory (speech) signals are represented bilaterally and symmetrically at an early sensory stage, and that the processing asymmetry emerges at a later stage due to asymmetric sampling in time (Poeppel, 2003; see also Zattore, Evans, Meyer, & Gjedde, 1992).

Ivry and Robertson (1998) further elaborated Sergent's hypothesis by proposing the Double Filtering by Frequency (DFF) theory, which posits that after attentional selection of a task-relevant frequency range, the LH amplifies high frequencies, whereas the RH amplifies low frequencies. Their model (Figure 2a) postulated different frequency tuning units and modules, and the output from each module was combined through an attentional weighting layer. They used one-dimensional hierarchical patterns (Figure 2b) in their simulations. Consistent with human data, the model exhibited a hemisphere-by-level interaction and a global level advantage (Figure 2c). Nevertheless, the underlying neural mechanism of this differential frequency filtering phenomenon remains unclear.

What could cause this asymmetry? One possibility is that there are anatomical differences between the hemispheres that influence processing. Recent research has shown that, in the left posterior superior temporal lobe, a region associated with language processing, pyramidal cells have longer dendrite lengths and contact fewer adjacent columnar units than do those in the RH (Anderson, Southern, & Powers, 1999; Buxhoeveden, Switala, Litaker, Roy, & Casanov, 2001; Hutsler & Galuske, 2003). A similar asymmetry also exists in the macrocolumnar structures (Galuske, Schlote, Bratzke, & Singer, 2000). In addition, Galuske et al. (2000) found that in the posterior part of Brodmann area 22, which involves language-relevant processing of auditory signals, there were modular networks of long-range intrinsic connections linking regularly spaced clusters of neurons; while the cluster size was similar in the two hemispheres, the spacing between clusters in the networks in the LH was about 20% larger than those in the RH. This asymmetry was not observed in the primary

auditory area. Although relevant anatomical data do not currently exist for the visual cortex, the behavioral asymmetry has been observed in both visual and auditory modalities (e.g., Ivry & Robertson, 1998; Hutsler & Galuske, 2003; Poeppel, 2003). We therefore hypothesize that there may be similar spacing differences in the left extrastriate areas versus the right.

Here we test the hypothesis that the perceptual asymmetry results from differential connection configurations at an encoding stage beyond the sensory level through computational modeling. In our model we use autoencoders (Figure 3), neural networks that learn compressed encodings of their input at the hidden layer (Rumelhart, Hinton, & Williams, 1986; Cottrell, Munro, & Zipser, 1987). The distribution of connections between the encodings and the input units is determined by a Gaussian probability density function (pdf). While holding the number of connections in each model fixed, we use a wide pdf to model longer-range connections between columns in the LH, and a narrow pdf to model short-range connections in the RH network. We then use a single-layer perceptron to extract from these encodings whether there is a target in the stimulus (either global or local). The error in the output reflects how informative the encoding is given the task, analogous to human reaction time – greater uncertainty leads to longer reaction times. Note here that the model's asymmetry is very different from the Gaussian receptive field functions used in previous models of hemispheric asymmetry. We sample from the Gaussian to allocate the occurrence of a fixed number of connections whose weights are set by learning, not as the activation function of a Radial Basis Function (RBF) unit (e.g., Monaghan & Shillcock, 2004), or as the weighting of the inputs (e.g., Ivry & Robertson, 1998). In fact, the receptive field widths in these prior models are the opposite of ours, i.e., wide in the RH, and narrow in the LH (e.g., Monaghan & Shillcock, 2004).

We conducted two simulations. In the first simulation, we used the same one-dimensional hierarchical pattern stimuli as the DFF simulation (Figure 2b; Ivry & Robertson, 1998). In the second simulation, we used hierarchical letter patterns

similar to those used in Sergent's experiment (Sergent, 1982; Figure 4); we also examined the resulting spatial frequency content after the differential encoding scheme was applied.

METHODS

Here we ran two types of simulations, both using two target patterns and two distracter patterns that could be combined into local targets and global distracters and vice-versa. In the first experiment, we used the simplified one-dimensional hierarchical stimuli used in Ivry and Robertson's (1998) simulation. Each stimulus was 29 units long, constructed by combining two patterns so that one pattern forms the local features and the other forms the global pattern of the stimulus, with a blank (0) unit between each local pattern (Figure 2b). In the second simulation, we replicated Sergent's experiment using two-dimensional hierarchical letter patterns. Each pattern could appear at the local or global level, for a total of 16 input patterns (Sergent, 1982). In this simulation, each pattern was 31 x 13 (403) pixels, with the same letters and same assignments of letters to targets and distracter sets as used in Sergent's experiment (Figure 4).

In the simulations, we used two autoencoder networks (Rumelhart et al., 1986; Cottrell, Munro, & Zipser, 1987) with different connectivity configurations as a way to learn an efficient encoding from the input data. While holding the number of connections for each hidden unit fixed, the LH network had a comparatively wider pattern of connectivity than the RH network (Figure 3), in accordance with the asymmetry reported between long-range connections in LH and RH BA22 (Galuske et al., 2000). More specifically, each hidden unit had a fixed number of connections to the input layer, and these connections were randomly drawn from a Gaussian probability density function (pdf). Each hidden unit within a model hemisphere used a Gaussian pdf with an identical σ (variance), with the LH σ ($\sigma_{1D} = 12$, $\sigma_{2D} = 18$; the subscripts 1D and 2D refer to the stimulations with one- and two-dimensional stimuli respectively) greater than the RH σ ($\sigma_{1D} = 1.8$, $\sigma_{2D} = 4$; see Figure 3). The variances

were chosen as two extreme cases of denseness/sparseness of the connections in order to examine the qualitative differences between the LH and RH networks; a wide range of values for the variances were tested, and similar results were found. The connection pattern from the hidden layer to the output layer was completely symmetric to those from the input layer to the hidden layer. Each hidden unit was associated with a position in the input space such that the set of hidden units were evenly distributed across the input space. When selecting the connections for a particular hidden unit, the Gaussian pdf was centered at that hidden unit's location in the input space.

After selecting all connections and constructing a network, the network was trained on all 16 input patterns until the network reached a fixed error (summed across all output units and patterns; see below for more details). Similarly to Monaghan and Shillcock (2004), we trained to a performance criterion, rather than for a fixed number of iterations, because the networks with different connectivity patterns learned the patterns at different rates. Once a network was trained, hidden unit encodings for each input pattern were computed by presenting the input pattern and then recording the hidden unit activities. These hidden unit encodings were compressed encodings that reflect the result of having differential connectivity to the hidden units.

After obtaining the compressed encodings of the input stimuli, we used a perceptron (i.e., a one-layer neural network) with a sigmoidal output function to classify the encodings according to whether there was a target or not (at either level) in the input stimuli, the same task subjects were required to do in Sergent's (1982) experiment. The output layer of the perceptron had a single node; the node had value '1' when a target was present at either level (75% of the stimuli) and '0' otherwise (25% of the stimuli). The error was measured as the difference between the output of the perceptron and the desired output (0 or 1). As has been done in previous studies, this error was considered to be a measure of uncertainty, and compared directly to human reaction time (e.g., Seidenberg & McClelland, 1989; Dailey, Cottrell, Padgett, & Adolphs, 2002).

In the simulation with one-dimensional stimuli, we explored the parameter space by testing the model with different combinations of the parameters, ranging from 11 to 15 hidden nodes and 5 to 10 connections from each hidden node. In the simulation with hierarchical letter patterns, the combinations ranged from 11 to 15 hidden nodes and 170 to 220 connections from each hidden node.

For both the autoencoder networks and the perceptron, the training algorithm was gradient descent (Rumelhart et al., 1986) using sum-squared error (SSE) for the objective function. The learning rate started at a constant ($\zeta_{1D} = \zeta_{2D} = 0.1$ for the autoencoder networks; $\zeta_{1D} = \zeta_{2D} = 0.05$ for the perceptron), and was adapted during training: if the error decreased in the current epoch, the learning rate for the next epoch increased by a factor of 1.05; if the error increased, the new learning rate was decreased by a factor of 1.25. Training of the autoencoders proceeded until the average SSE across all output nodes reached a pre-determined threshold (0.025) within a pre-determined maximum number of iterations ($max_{1D}=1000$, $max_{2D}=250$). Rare cases where the autoencoder could not reach the SSE performance criterion within the maximum number of training iterations were marked as rejections. Little effect was seen in varying this threshold in the ranges of 0.05 (requiring very few training iterations) to 0.01 (requiring many training iterations and leading to a high incidence of rejections). Training for the perceptron classifiers stopped after 250 iterations; values between 100 and 1000 iterations showed similar performance. After training the perceptron had 100% classification accuracy.

In order to match the statistical power found in Sergent's experiment, we ran the model 68 times in each simulation, giving us approximately the same number of total trials (68 models x 16 trials per model hemisphere) as Sergent's human data (12 subjects x 90 trials per visual field).

In order to examine encoding differences between LH and RH networks in terms of spatial frequency, output images were computed for each network. This was done by presenting each input image to a trained network, then recording the output unit activities. These output images were then analyzed for spatial frequency content.

In order to compare and visualize, we took the log-power at each frequency, then computed the difference in log-power between RH and LH networks. We used hierarchical letter patterns (Sergent, 1982) for this analysis.

RESULTS

Results of the simulation with one-dimensional stimuli

We first report the results of the simulation in which we used the same one-dimensional stimuli as those used in Ivry and Robertson's (1998) model (Figure 2b). In order to verify that the results were robust to the parameters defining the model architecture, we ran the model with different parameter combinations, ranging from 11 to 15 hidden nodes and 5 to 9 connections from each hidden node (in total 25 different combinations). We used repeated measures ANOVA to analyze the data; the within-subject variable was target level (global vs. local), whereas the between-subject variables were hemisphere (LH vs. RH networks), number of hidden nodes (11, 12, 13, 14, and 15), and number of connections from each hidden node (5, 6, 7, 8, and 9). The dependent variable was the error in the output layer of the perceptron.

Consistent with human data, the results showed that the model had better performance when the target was at the global level ($F(1, 3350) = 1092.823$, $p < 0.001$), and there was a significant interaction between hemisphere and target level ($F(1, 3350) = 756.923$, $p < 0.001$; Figure 5a); although both of these two effects interacted with either the number of hidden nodes (target level x number of hidden nodes, $F(4, 3350) = 38.347$, $p < 0.001$; target level x hemisphere x number of hidden nodes, $F(4, 3350) = 20.456$, $p < 0.001$), or number of connections (target level x number of connections, $F(4, 3350) = 11.927$, $p < 0.001$; target level x hemisphere x number of connections, $F(4, 3350) = 4.805$, $p = 0.001$), when we split the data according to either number of hidden nodes or number of connections, both effects were significant in all cases ($p < 0.001$ for all cases; Figure 5b). Nevertheless, in Sergent's (1982) human data, there was no main effect of hemisphere; the two hemispheres had a similar performance level on average. In contrast, our model

showed a main effect of hemisphere: the LH network performed better than the RH network ($F(1, 3350) = 154.231$, $p < 0.001$); this effect interacted with number of hidden nodes ($F(1, 3350) = 12.808$, $P < 0.001$): performance difference between the two hemisphere networks was significant when the network had 11 ($F(1, 670) = 69.770$, $p < 0.001$), 12 ($F(1, 670) = 63.882$, $p < 0.001$), 13 ($F(1, 670) = 16.954$, $p < 0.001$), or 14 hidden nodes ($F(1, 670) = 19.119$, $p < 0.001$), but not when it had 15 hidden nodes ($F(1, 670) = 2.383$, $p = 0.123$; Figure 5b). This suggests that performance difference between the two hemisphere networks can be influenced by parameter settings.

Results of the simulation with hierarchical letter pattern stimuli

In the second simulation, we used the hierarchical letter patterns used in Sergent's study. We explored how the performance changed with different parameter combinations, ranging from 22 to 30 hidden nodes and 40 to 120 connections from each hidden node (in total 25 different combinations). As in the first simulation, we used repeated measures ANOVA to analyze the data; the within-subject variable was target level (global vs. local), whereas the between-subject variables were hemisphere (LH vs. RH networks), number of hidden nodes (22, 24, 26, 28, and 30), and number of connections from each hidden node (40, 60, 80, 100, and 120). The dependent variable was the error in the output layer of the perceptron.

The results showed an advantage of detecting a global level target ($F(1, 3350) = 1070.838$, $p < 0.001$) and an interaction between hemisphere and target level ($F(1, 3350) = 858.284$, $p < 0.001$; Figure 6a); both effects interacted with number of connections ($F(4, 3350) = 36.261$, $p < 0.001$) but not number of hidden nodes ($F(4, 3350) = 1.933$, $p = 0.102$). When we split the data by number of connections, we found that both effects were significant across all cases (Figure 6b). The model also showed a main effect of hemisphere ($F(1, 3350) = 91.054$, $p < 0.001$), and this effect interacted with number of connections ($F(4, 3350) = 17.519$, $p < 0.001$): when the model had 40 ($F(1, 670) = 58.135$, $p < 0.001$), 60 ($F(1, 670) = 112.781$, $p < 0.001$), or

80 connections from each hidden node ($F(1, 670) = 14.713$, $p < 0.001$), the LH network performed significantly better than the RH network; this difference was not significant when the model had 100 ($F(1, 670) = 1.343$, $p = 0.247$) or 120 connections ($F(1, 670) = 0.251$, $p = 0.617$).

Thus, the results from the two simulations suggested that although the global level advantage effect and the interaction between hemisphere and target level could be modulated by different parameter settings, the modulation generally only affected the size of the effects, not the direction; in other words, these effects were robust against parameter changes. In contrast, the performance difference between the LH and RH networks was sensitive to parameter settings[1].

We also investigated spatial frequency content preserved in the LH and RH encodings. We reproduced input images from their encodings in the output; for hierarchical letter patterns, low frequencies were better reproduced in the RH network, whereas high frequencies were better reproduced in the LH network (Figure 7a and 7b, consistent with Sergent's (1982) hypothesis and the DFF theory (Ivry & Robertson, 1998). However, this did not result directly from frequency tuning of the neurons. Rather, differential frequency filtering behavior emerged naturally as the result of the encoding scheme, suggesting that the asymmetry in perception may be due to differences in anatomy rather than frequency tuning per se.

DISCUSSION

In the current study, we test the hypothesis that hemispheric asymmetry in the perception of global and local features originates from differential encoding beyond the sensory level due to anatomical differences between the two hemispheres, instead of differential frequency filtering as proposed by the DFF theory (Ivry & Robertson, 1998). We first argue that the lack of evidence supporting hemispheric specialization for particular frequency ranges (e.g., Kitterle et al., 1990; Rijsdijk et al., 1980; Di Lollo, 1981; Peterzell et al., 1989; Fendrich & Gazzaniga, 1990) suggests that this hemispheric asymmetry takes place beyond the sensory level (Sergent, 1982; Heinze,

et al., 1998) and the two hemispheres do not differ in information extraction. We then argue that the difference takes place at an encoding stage due to differences in connection structures. We incorporate evidence about the anatomical differences in columnar and connectional structure in the auditory cortex between the two hemispheres (e.g., Galuske et al., 2000; Seldon, 1981a; 1981b; 1982; Anderson, Southern, & Powers, 1999; Buxhoeveden, Switala, Litaker, Roy, & Casanova, 2001; Hutsler & Galuske , 2003) into a computational model that uses autoencoder networks to develop efficient encodings of the stimuli (Rumelhart et al., 1986; Cottrell, Munro & Zipser, 1987): the columnar structure in the posterior superior temporal lobe in the RH has more connections among neighboring columns compared with the LH, and thus may develop representations that are more functionally overlapped than those in the LH (Hutsler & Galuske, 2003). Although relevant anatomical data for the visual cortex are not currently available, similar perceptual asymmetry has been observed in both visual and auditory modalities (e.g., Poeppel, 2003). Thus, based on a hypothesized generalization across the two modalities, we use two autoencoder networks with differential connectivity configurations to simulate this differential encoding: the RH autoencoder network has a narrower connection distribution to allow more connections among neighboring nodes compared with the LH autoencoder network. We then use a perceptron to examine how efficacious the two encoding systems are in terms of detecting local and global level targets. The results match human data (Sergent, 1982) well; they show a significant hemisphere-by-level interaction: a RH advantage for responses to a global level target, and a LH advantage for responses to a local level target (Sergent, 1982). They also show an overall advantage in responses to a global level target, consistent with human data (Navon, 1977). This effect is because the narrower connection distribution in the RH autoencoder network allows each hidden node to develop a compressed representation for a local region within the stimulus; since in natural images neighboring pixels are more correlated than distant ones, there may be more variance in low spatial frequencies across the input patterns received by a hidden node, resulting in the

dominance of low spatial frequency information. In contrast, with a wider and sparser connection distribution, each hidden node in the LH autoencoder network samples across a wider range of the input image and the sampled pixels are more random and less likely to be correlated; consequently there may be comparable variance in high and low spatial frequencies across the input patterns received by a hidden node, resulting in the LH network's better ability in preserving high spatial frequencies as compared with the RH network[2].

In comparison with Ivry and Robertson's (1998) DFF model, we show that our model provides a better account of human data (Sergent, 1982). Their model enforces a discrete separation of frequency information into modules, and hemispheric differences take place through manipulating the combination of the outputs from different frequency modules. It is unclear how these frequency ranges are combined in a certain way, and how the model is able to account for the lack of evidence supporting hemispheric specialization for particular frequency ranges (Kitterle et al., 1990; Rijsdijk et al., 1980; Di Lollo, 1981; Peterzell et al., 1989; Fendrich & Gazzaniga, 1990). In addition, there is little anatomical evidence suggesting differential frequency tuning in the neurons in the two hemispheres, or differential modulation by frequency channels in the two hemispheres similar to that proposed in the DFF model. In contrast, through hypothesizing that hemispheric differences take place at an encoding stage beyond the sensory level, and using Gaussian probability distributions to simulate differential connection configurations at the encoding stage, our model naturally develops the hemispheric difference in the frequency content in the encoding.

In our simulation with one-dimensional stimuli as those used in the DFF model, we explored the parameter space and found that the main effect of global level advantage and the interaction between network and target level were robust against parameter changes, although in some cases there was a significant main effect of LH network advantage. In contrast, in the DFF model, with one given parameter setting, the interaction between network and target level was fragile – the LH network became

better at identifying both local and global targets with further training. Also, the simulation of the DFF model used one-dimensional hierarchical stimuli that differed greatly from Sergent's original hierarchical letter patterns. In contrast, here we used two-dimensional hierarchical letter patterns similar to those used in human studies (Sergent, 1982) and replicated the results, a test that has not been conducted with the DFF model. In addition, through analyzing the spatial frequency content preserved in the encodings from the LH and RH networks, we show that differential frequency filtering behavior emerged naturally as the result of the encoding scheme, suggesting that hemispheric asymmetry in perception may be due to hemispheric differences in connection structures rather than frequency tuning per se.

The modeling results provide support for the idea that a hemispheric difference in cortical columnar and connection structure similar to that in the auditory cortex may also exist in high-level visual areas. We speculate that it may be in the lateral occipital region. It has been reported that there is significantly greater ipsilateral activity (i.e., activation from the other hemisphere after the initial contralateral projection from the visual hemifields to the hemispheres) observed in the area anterior to the retinotopic areas (Tootell, Mendola, Hadjikhani, Liu, & Dale, 1998), suggesting that the lateral occipital region may be a convergence point after the visual field split (Hsiao, Shieh, & Cottrell, 2008). Consistent with this speculation, recent fMRI studies have suggested that the locus of this hemispheric asymmetry in local and global processing is in the occipital/occipitotemporal region (Martinez et al., 1997; Han et al., 2002). Another possibility is the inferior parietal lobe/superior temporal gyrus region, suggested by recent fMRI studies showing that the activation in this region corresponds to the asymmetry observed in human data (Weissman & Woldorff, 2005; Robertson, Lamb, & Knight, 1988; Fink et al., 1997). Further examinations are required to confirm these speculations.

We are currently pursuing the incorporation of more anatomical data into the model, such as using 2D Gabor filters to simulate responses of complex cells in the early visual system (Daugman, 1985), and also using the proposed autoencoder

networks as the way to develop efficient encoding in the two hemispheres in modeling more complicated real world visual stimuli (such as faces; cf., the Principal Component Analysis step in many visual perception models, e.g., Hsiao et al., 2008; Dailey & Cottrell, 1999; Dailey, Cottrell, Padgett, & Adolphs, 2002), in order to further examine the cognitive plausibility of this differential encoding mechanism in accounting for other hemispheric asymmetry phenomena in perception, such as the left side bias in face perception (e.g., Gilbert & Bakan, 1973) and the RVF advantage in visual word recognition (e.g., Bryden & Rainey, 1963).

REFERENCES

Anderson, B., Southern, B. D., & Powers, R. E. (1999). Anatomic asymmetries of the posterior superior temporal lobes: a postmortem study. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*, *12*, 247–254.

Bartholomeus, B. (1974). Effects of task requirements on ear superiority for sung speech. *Cortex*, *10*, 215-223.

Bryden, M.P., & Rainey, C.A. (1963). Left-right differences in tachistoscopic recognition. *Journal of Experimental Psychology*, *66*, 568–571.

Buxhoeveden, D. P., Switala, A. E., Litaker, M., Roy, E., & Casanova, M. F. (2001). Lateralization of minicolumns in human planum temporal is absent in nonhuman primate cortex. *Brain, Behavior, & Evolution*, *57*, 349-358.

Chan, K. W., & Hsiao, J. H. (in press). Hemispheric asymmetry in processing low- and high-pass filtered Cantonese speech in tonal and non-tonal language speakers. *Language & Cognitive Processes*.

Cottrell, G., Munro, P. and Zipser, D. (1987) Learning internal representations from gray-scale images: An example of extensional programming. *Proceedings of the*

Ninth Annual Cognitive Science Society Conference, Seattle, WA, July 16-18. (pp. 461-473)

Dailey, M. N., & Cottrell, G. W. (1999). Organization of Face and Object Recognition in Modular Neural Networks. *Neural Networks*, *12(7-8)*, 1053-1074.

Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (2002). EMPATH: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience*, 14, 1158-1173.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, *2*, 1160-1169.

Delis, D. C., Robertson, L. C., & Efron, R. (1986). Hemispheric specialization of memory for visual hierarchical stimuli. *Neuropsychologia*, *24*, 205–214.

Di Lollo, V. (1981). Hemispheric symmetry in visible persistence. *Perception & Psychophysics*, *11*, 139-142.

Fendrich, R., & Gazzaniga, M. (1990). Hemispheric processing of spatial frequencies in two commissurotomy patients. *Neuropsychologia*, *28,* 657–663.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain*, *120*, 1779–1791.

Flevaris, A. V., Bentin, S., & Robertson, L. C. (2010). Local or global? Attentional selection of spatial frequencies binds shapes to hierarchical levels. *Psychological Science*, *21*, 424-431.

Flevaris, A. V., Bentin, S., & Robertson, L. C. (2011). SF mediates global versus local processing: Evidence from EEG. *Journal of Vision*, *11(7)*, 11.

Galuske, R. A., Schlote, W., Bratzke, H., & Singer, W. (2000). Interhemispheric asymmetries of the modular structure in human temporal cortex. *Science*, *289*, 1946-1949.

Gilbert, C., & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, *11*, 355-362.

Han, S., Weaver, J. A, Murray, S. O., Kang, X, Yund, E. W. & Woods, D. L. (2002). Hemispheric asymmetry in global/local processing: effects of stimulus position and spatial frequency. *Neuroimage*, *17*, 1290-1299.

Heinze, H. J., Hinrichs, H., Scholz, M., Burchert, W., & Mangun, G. R. (1998). Neural mechanisms of global and local processing: A combined PET and ERP study. *Journal of Cognitive Neuroscience*, *10*, 485–498.

Hoffman, J. E. (1980). Interaction between global and local levels of a form. *Journal of Experimental Psychology: Human Perception & Performance*, *6*, 222-234.

Hsiao, J. H., Shieh, D., & Cottrell, G. W. (2008). Convergence of the visual field split: hemispheric modeling of face and object recognition. *Journal of Cognitive Neuroscience*, *20*, 2298-2307.

Hutsler, J., & Galuske, R. A. W. (2003). Hemispheric asymmetries in cerebral cortical networks. *Trends in Neurosciences*, *26*, 429-435.

Ivry, R., & Lebby, P. (1993). Hemispheric differences in auditory perception are similar those found in visual perception. *Psychological Science*, *4*, 41-45.

Ivry, R., & Robertson, L. C. (1998). *The Two Sides of Perception*. Cambridge: MIT Press.

Keenan, P. A., Whitman, P. D., & Pepe, J. (1989). Hemispheric asymmetry in the processing of high and low spatial frequencies: A facial recognition task. *Brain & Cognition* , *11*, 229-237.

Kitterle, F. L., Christman, S., & Hellige, J. B. (1990). Hemispheric differences are found in the identification, but not the detection, of low versus high spatial frequencies. *Perception & Psychophysics*, *48*, 297-306.

Ley, R. G., & Bryden, M. P. (1982). A dissociation of right and left hemisphere effects for recognizing emotional tone and verbal content. *Brain & Cognition*, *1*, 3-9.

Martin, M. (1979). Hemispheric specialization for local and global processing. *Neuropsychologia*, *17*, 33-40.

Martinez, A., Moses, P., Frank, L., Buxton, R., Wong, E., & Stiles, J. (1997). Hemispheric asymmetries in global and local processing: Evidence from fMRI. *NeuroReport*, *8*, 1685–1689.

Monaghan, P., & Shillcock, R. C. (2004). Hemispheric asymmetries in cognitive modeling: connectionist modeling of unilateral visual neglect. *Psychological Review*, *111*, 283-308.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353-383.

Peterzell, D. H. (1991). On the nonrelationship between spatial frequency and cerebral hemispheric competence. *Brain & Cognition*, *15,* 62-68.

Peterzell, D. H., Harvey, L. O., Jr., & Hardyck, C. D. (1989). Spatial Frequencies and the cerebral hemispheres: Contrast sensitivity, visible persistence, and letter classification. *Perception & Psychophysics*, *46*, 433-455.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as asymmetric sampling in time. *Speech Communication*, *41*, 245-255.

Proverbio, A. M., Minniti, A., & Zani, A. (1998). Electrophysiological evidence of a
 perceptual precedence of global vs. local visual information. *Cognitive Brain
 Research*, *6*, 321-334.

Proverbio, A. M., Zani, A., & Avella, C. (2002). Hemispheric Asymmetries for
 Spatial Frequency Discrimination in a Selective Attention Task. *Brain &
 Cognition*, *34*, 311-320.

Rijsdijk, J. P., Kroon, J. N., &Van der Wildt, G. J. (1980). Contrast sensitivity as a
 function of position on retina. *Vision Research* , *20*, 235-241.

Robertson, L. C., & Delis, D. C. (1986). 'Part-whole' processing in unilateral brain-
 damaged patients: Dysfunction of hierarchical organization. *Neuropsychologia*, *24*,
 363–370.

Robertson, L. C., Lamb, M. R., & Knight, R. T. (1988). Effects of lesions of
 temporoparietal junction on perceptual and attentional processing in humans.
 Journal of Neuroscience, *8*, 3757-3769.

Robertson, L. C., Lamb, M. R., & Zaidel, E. (1993). Interhemispheric relations in
 processing hierarchical patterns: Evidence from normal and commissurotomized
 subjects. *Neuropsychology*, *7*, 325–342.

Rumelhart, D. E., Hinton, G. E, & Williams, R. J. (1986). Learning representations by
 back-propagating errors. *Nature*, *323*, 533-536.

Seidenberg, M. S., & McClelland, J. L. (1989). A Distributed, Developmental Model
 of Word Recognition and Naming. *Psychological Review, 96,* 523-568.

Seldon, H. L. (1981). Structure of human auditory cortex. I. Cytoarchitectonics and
 dendritic distributions. *Brain Research*, *229*, 277–294.

Seldon, H. L. (1981). Structure of human auditory cortex. II. Axon distributions and
 morphological correlates of speech perception. *Brain Research*, *229*, 295–310.

Seldon, H. L. (1982). Structure of human auditory cortex. III. Statistical analysis of

    dendritic trees. *Brain Research*, *249*, 211–221.

Sergent, J. (1982). The Cerebral Balance of Power: Confrontation or Cooperation?

    Journal of Experimental Psychology: Human Perception & Performance, *8*, 253-

    272.

Tootell, R. B. H., Mendola, J. D., Hadjikhani, N. K., Liu, A. K., & Dale, A. M. (1998).

    The representation of the ipsilateral visual field in human cerebral cortex.

    Proceedings of the National Academy of Sciences USA, *95*, 818–824.

Van Kleeck, M. H. (1989). Hemispheric differences in global versus local processing

    of hierarchical visual stimuli by normal subjects: New data and a meta-analysis of

    previous studies. *Neuropsychologia*, 27, 1165–1178.

Weissman, D. H., & Woldorff, M. G. (2005). Hemispheric asymmetries for different

    components of global/local attention occur in distinct temporo-parietal loci.

    Cerebral Cortex, *15*, 870-876.

Zattore, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of

    phonetic and pitch discrimination in speech processing. *Science*, *256*, 846-849.

Author Notes

Janet H. Hsiao, Department of Psychology, University of Hong Kong, Hong Kong.

Ben Cipollini, Department of Cognitive Science, University of California, San Diego, USA.

Garrison W. Cottrell, Department of Computer Science & Engineering, University of California, San Diego, USA.

Correspondence concerning this article should be addressed to Janet H. Hsiao, Department of Psychology, University of Hong Kong, Pofuklam Road, Hong Kong, or to the authors at jhsiao@hku.hk, bcipolli@cogsci.ucsd.edu, or gary@cs.ucsd.edu, respectively.

Footnotes

[1] In our simulations, we consistently found a significant interaction between target level and hemisphere, a significant LH advantage over the RH when the target was at the local level, and a significant RH advantage over the LH when the target was at the global level, consistent with the human data. However, when we examined the data of the LH and RH networks separately, whereas a strong global level advantage over the local level condition was observed in the RH network, there was no apparent local level advantage in the LH network (Figure 5 and 6). In Sergent's (1982) results (Figure 1), the difference between L+S- and L-S+ conditions in the RVF/LH presentation condition was also much smaller than that in the LVF/RH presentation condition; whether this difference was significant was not reported.

[2] In a separate simulation, we used low-pass and high-pass filtered hierarchical patterns as the stimuli. We found that the RH network had better performance in reproducing low-pass filtered stimuli than the LH network, whereas the LH network had better performance in reproducing high-pass filtered stimuli. This result further confirms that the RH network is biased to learn and represent low spatial frequency information, and the LH network is biased to learn and represent high spatial frequency information.

Figure 1. (a) Stimuli in Sergent's (1982) experiment. A hierarchical letter pattern contains a global and a local pattern; the global pattern (the large letter) is composed of a number of local patterns (the small letters). Sergent used four letters to compose the patterns: "H" and "L" were designated as targets, and "T" and "F" as distracters. "L+" means the large letter is a target, and "S+" means the small letters are targets. "id." means the local and global patterns are identical. (b) The RT data for the L+S- and L-S+ stimuli in the LVF and RVF presentation conditions (Sergent, 1982).
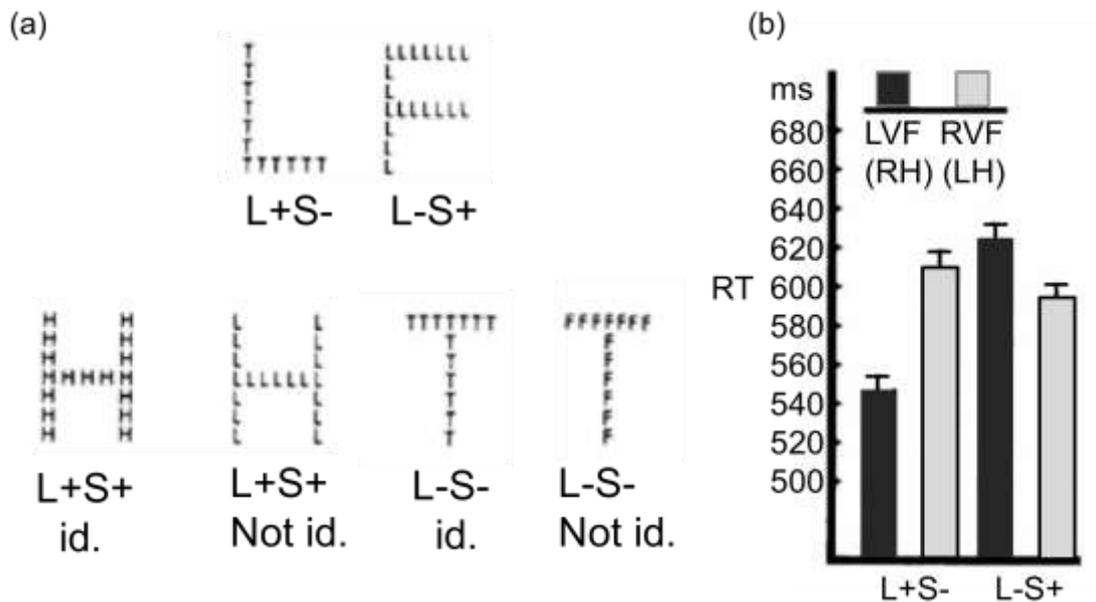
Figure 2. (a) Ivry and Robertson's computational model based on the DFF theory (Ivry & Robertson, 1998). The model contains six different frequency modules; each module extracts information of a specific spatial frequency from the input and learns to map it to the output (Module 6 has the lowest frequency). The four decision nodes correspond to four target patterns: whether target 1 or 2 is present, and whether it is at the global or local level. The outputs from the modules then go through an *attention weight* layer as a filter. The filter first selects a task-relevant frequency range; at the second stage, in the RH network, it amplifies the output from the low spatial frequency modules within the range, whereas in the LH network it amplifies the output from the high spatial frequency modules, through giving different weights to different modules. The figure shows an example RH network. (b) One-dimensional hierarchical patterns. There are two target (10101 and 01110) and two distracter patterns (11010 and 10110). Shown at the top is an actual input pattern formed by taking the first distracter pattern and replacing each black portion with a target pattern; this represents the first target pattern at the local level and the second distracter pattern at the global level. A 0 unit appears between each local pattern as a separator. (c) Results of the model with large stimuli (i.e., stimuli are enlarged by five) after 100 epochs showed an advantage for stimuli with a global level target and an interaction between network and target level, consistent with human data. Note that the LH network became better at identifying both local and global targets with further training.
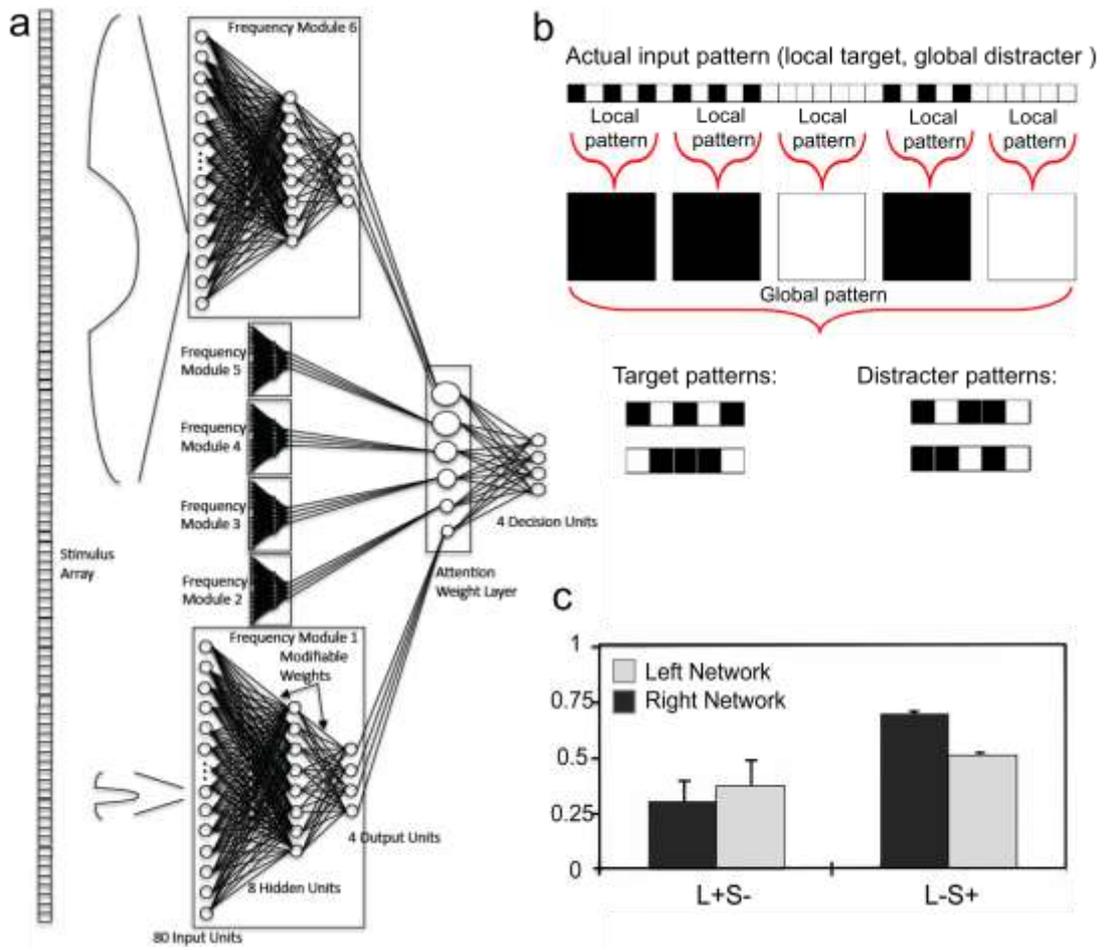
a

Frequency Module 6

Frequency Module 5

Frequency Module 4

Frequency Module 3

Frequency Module 2

Frequency Module 1

Modifiable Weights

4 Decision Units

Attention Weight Layer

Stimulus Array

4 Output Units

8 Hidden Units

80 Input Units

b

Actual input pattern (local target, global distracter)

Local pattern Local pattern Local pattern Local pattern Local pattern

Global pattern

Target patterns:

Distracter patterns:

c

Left Network
Right Network

1
0.75
0.5
0.25
0

L+S-   L-S+

26

Figure 3. LH and RH autoencoder networks; both have the same number of connections. Each hidden node has a fixed number of symmetric connections to the input and output layers respectively.
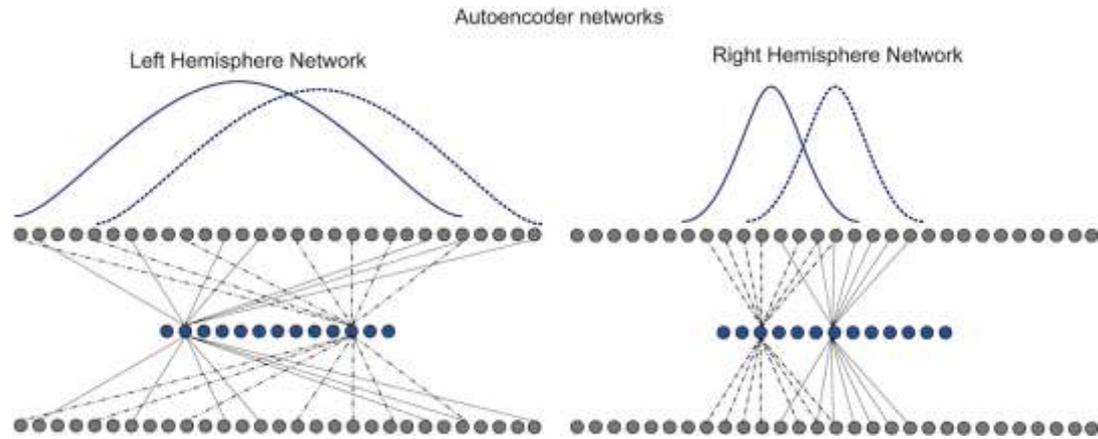

Autoencoder networks

Figure 4. Hierarchical letter patterns used in our second simulation. Each pattern is 31 X 13 (403) pixels. They are composed of the same letters used in Sergent's (1982) experiment.
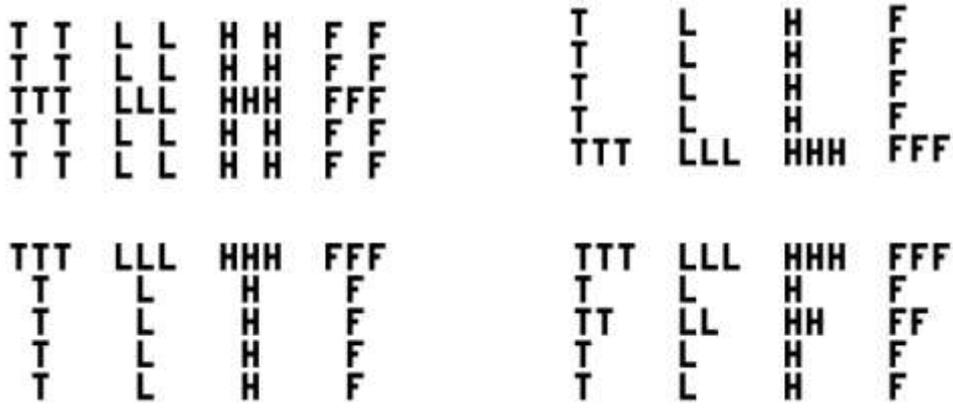
Figure 5. (a) Results of the simulation with one-dimensional stimuli used in the DFF model (Ivry & Robertson, 1998). (b) Results of the simulation when splitting the data according to either number of hidden nodes, or number of connections.
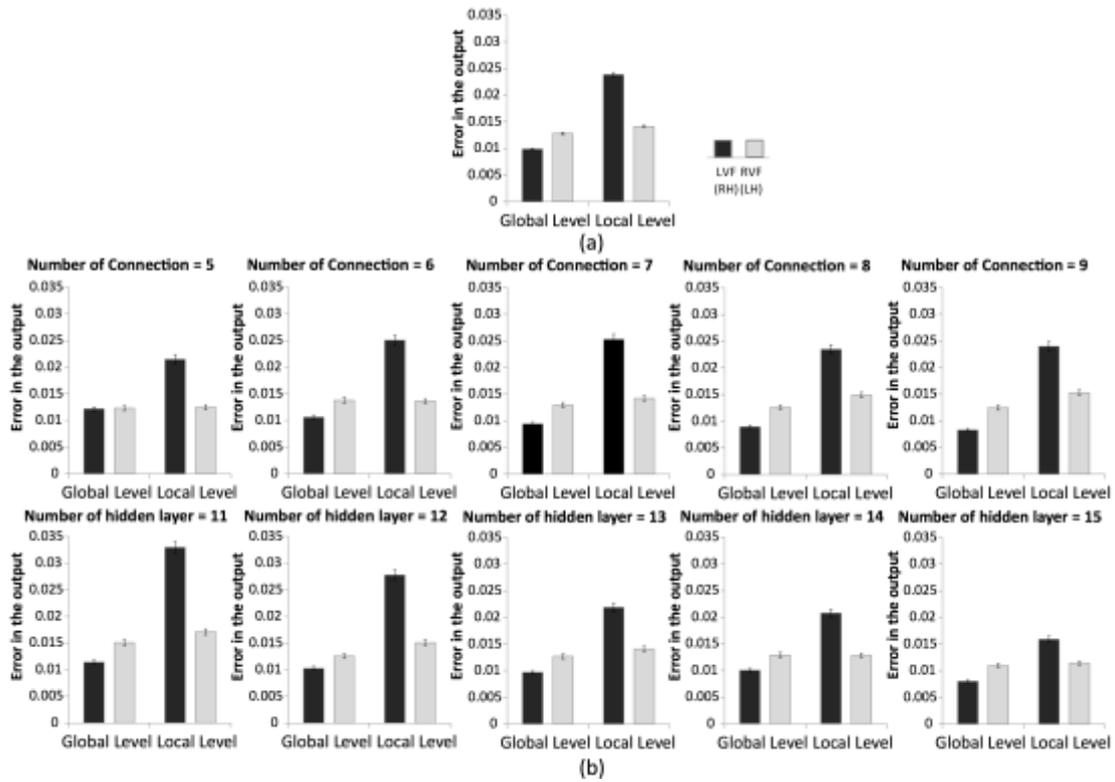
Figure 6. (a) Results of the simulation with hierarchical letter pattern stimuli. (b) Results of the simulation when splitting the data according to number of connections from each hidden node.
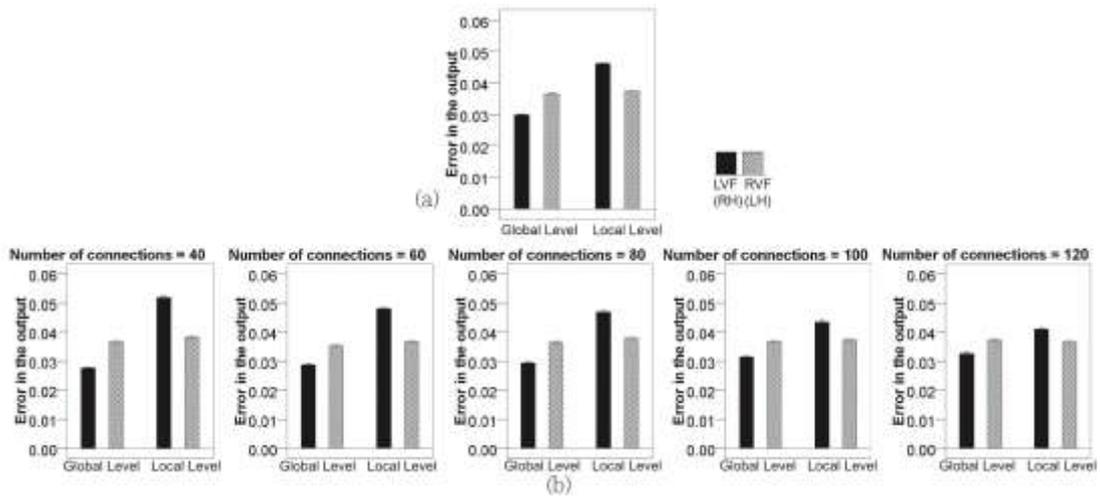
Figure 7. (a) Image reproduction example (Global: H; local: F) showing the frequency information in which the two networks significantly differ in power, and (b) Spatial frequency analysis of the output from the autoencoders with 26 hidden nodes and 100 connections to/from each hidden node in the simulation with hierarchical letter pattern stimuli. The plots show the difference in log radially averaged power spectrum (i.e., the directional independent mean spectrum) between the two networks (RH – LH); the blue line shows the mean whereas the red dash line indicates one standard deviation across the 68 simulation runs. Regions marked in yellow indicate significant difference from zero.