

Patchy Connectivity and Visual Processing Asymmetries: A Neurodevelopmental Hypothesis

Ben Cipollini

UC San Diego

Garrison W. Cottrell

UC San Diego

The existence of left-right asymmetries in classification of visual stimuli has been established by a long history of behavioral experiments and neuroimaging studies. The Double Filtering by Frequency (DFF) model¹¹ accounts for experimental data from four key publications, but offers no neural basis for a spatial frequency processing asymmetry it postulates to exist. The Differential Encoding (DE) model¹⁰ is neurally inspired, but has accounted for very little experimental data and hasn't followed the underlying biology closely. In this paper, we address the issues with the DE model in three steps. First, we make a more direct mapping between the DE model and the connectivity asymmetry that inspired it. Second, we apply the model to a similar set of data accounted for by the DFF model. Third, we show that the proposed connectivity asymmetry can emerge through typical development, and corroborate the relationship between connectivity and spatial frequency processing. Overall, we find that this version of the DE model can account for three of the four datasets tested: processing of local/global levels in Navon figures¹⁷ and using high/low frequency information based on task demands¹³, and a dataset showing lateralization of face identification²¹. The final dataset is from an experiment used to demonstrate that it is relative frequency in the stimuli, not absolute frequency differences, that matter. We suggest that a deeper version of the model could account for the data in a neurally plausible way, using absolute frequency differences encountered at different stages of processing.

Keywords: Asymmetry; Differential Encoding; Double filtering by Frequency; Hierarchical Letter Stimuli; Lateralization; Visual Processing.

1. Introduction

A large literature of behavioral experiments and neuroimaging studies has established the existence of left-right asymmetries in classification tasks across a variety of visual stimuli. Typically, these studies have used visual half-field presentation, where a stimulus is briefly presented to the left or right of a fixation point, then subjects press a button to indicate their response to some type of classification task (such as whether a target stimulus was present or not). If the same stimulus is presented in the right visual field (RVF) and left visual field (LVF) and

differences in reaction time or accuracy are found, then a hemispheric difference in the processing of the stimulus is inferred due to constrained connectivity in the brain, where information from the LVF is initially directed exclusively to the right cerebral hemisphere (RH), and information from the RVF to the left cerebral hemisphere (LH). Across these studies, the two inferences that are generally made are that the LH shows a specialization for processing information at higher spatial frequencies (HSF), while the RH shows a specialization for processing information at lower spatial frequencies (LSF)^{7,11,17}.

Ivry and Robertson¹¹ developed the Double Filtering by Frequency (DFF) theory to account for these visual half-field results as well as asymmetries in frequency processing of auditory stimuli. They proposed that, after early sensory encoding decomposes a stimulus into its component frequencies, processing involves two filtering stages. First, each hemisphere focuses on only the frequencies relevant to the task. Then, each hemisphere has a bias in encoding information within that band of task-relevant frequencies: the LH encodes more information in the higher frequencies of those selected in the first stage, while the RH encodes more spectral power in the lower frequencies of those selected.

To provide some validation of their theory, Ivry and Robertson also implemented a neural network model based on the DFF theory. This model was able to account for three particular experiments from the literature, each of which Ivry and Robertson argued expresses a core feature of visual processing asymmetries. These include a study of local/global processing, and two studies examining classification of spatial frequency gratings. Later, Hsiao *et al.*⁹ implemented a different neural network model of the DFF theory, and showed that it could account for a fourth study on face processing (see Table 1 for more information on these studies).

While the DFF theory and associated models account for all of these data, Hsiao *et al.*¹⁰ questioned some of the assumptions and implementation details of these models, including:

- There is a lack of physiological evidence suggesting differential frequency tuning in the neurons of the two hemispheres.
- The stimuli were drastically reduced in complexity relative to the experimental stimuli: they used 1D arrays (instead of 2D images) with only 2 stimulus classes (instead of the 6 presented to human subjects).
- The classification task used to train their model differed from that given to human subjects.
- With further training, their hemisphere \times level interaction disappeared or reversed.

To address this, Hsiao *et al.*¹⁰ developed a neural network model with a simple architectural difference between the hemispheres. Specifically, they incorporated an anatomical asymmetry in long-range lateral connections between “patches” found in Wernicke’s area (a language-related auditory area) and its RH homologue⁴ into a simple autoencoder neural network. These connections selectively interconnect a subset of nearby patches, or macro columns, and are involved in combining information across classic receptive fields. This Differential Encoding (DE) model was used to encode 2D images of Navon figures (large “global”-level letters composed of smaller “local”-level letters); when these encodings were then classified with a second neural network, the same hemisphere \times level interaction emerged.

While suggestive, this model was used to account for only one of the four experiments that the DFF has accounted for. In addition, the DE model did not use biologically plausible parameters, nor was the specific role of the connections in question justified in the experiment modeled. Finally, how or why this connectivity difference might occur during development was never addressed.

We aim to address all of these issues here. We begin by arguing that long-range lateral connections are likely to be involved in the experimental paradigms found to elicit visual processing asymmetries, due to their activation when stimulus strength is low. We argue that the DE model, with a specific configuration, is a reasonable implementation of the anatomical system. We then describe modeling experiments where we implemented this model setup, and showed that such a model exhibits both spatial frequency differences, as well as differences on classification tasks of encoded stimuli, for three of the behavioral studies that the DFF has accounted for. We discuss why we think a failure to account for the fourth study is informative and potentially interesting. Finally, to support the plausibility of our work, we show, using a developmental version of our model, that normal neural development of visual cortex, in conjunction with changes in visual acuity and a hypothesized timing difference in the maturation of the hemispheres, can lead to the hypothesized anatomical asymmetry.

2. Patchy connectivity may be involved in processing asymmetries

Patches are groups of locally interconnected mini-columns, spanning about 500 μ m in diameter in cortex. They are called patches because an injection of dye into the cortical surface will label cortex at the injection site, as well as a number of discrete patches of surrounding cortex. They are thought to be functional units, and somewhat akin to hypercolumns. These patches themselves interconnect with nearby patches within a cortical area through long-range lateral connections that

travel through the grey matter. These patchy connection networks are quite specific and sparse; a patch will connect with a subset of its local neighbors¹⁴.

Why might long-range lateral connections be specifically involved in Sergent’s experiment, or other experiments showing visual processing asymmetries and accounted for by Ivry and Robertson’s DFF model? Looking into the literature, we found a number of reviews showing that asymmetries were quite dependent on task setup and stimulus parameters, specifically stimulus contrast and stimulus presentation duration: stimuli must be presented with low contrast, or briefly, to elicit asymmetries^{2,11,18,19}. When the stimulus strength is high, asymmetries disappear.

We interpret these findings to suggest that asymmetry is not in the feed-forward pathway, whose function we assume is positively correlated to stimulus strength, but instead can be found in secondary types of connections - feed-back and long-range lateral connections. We focus on long-range lateral connections between patches because they have specifically been found to be more active during low stimulus strength²⁰, are well characterized anatomically and physiologically¹⁴, are thought to help encode higher-level visual information such as contours^{5,6}, and as described above, we know of at least one existing asymmetry in this type of connectivity⁴.

We find that a specific implementation of the neural network employed by Hsiao *et al.* to be a plausible model of how long-range lateral connections affect a map-based encoding, such as a retinotopic encoding of an image. Specifically, connections from the input to hidden layer can be interpreted as a recurrent network unfolded in time, if each hidden unit connects to the input image at that hidden unit’s location. In this particular model architecture, forward propagation of activity along the input-to-hidden unit connections would represent the propagation of activity laterally, over one time-step. The number of hidden units used would represent the number of local patchy connection networks used to reconstruct the image; when the number of hidden units matches the number of pixels, then all patchy connection networks are used in the reconstruction.

2.1. Methods

The Differential Encoding (DE) model consists of two autoencoder neural networks, each representing one cerebral hemisphere, which differ only in the spatial spread of connections from the input to hidden, and hidden to output layers. Unlike most autoencoders, the hidden units of these models connect to a small subset of the input and output banks (see Figure 1). Each hidden unit has a position in the input (and output) arrays, and a fixed number of connections to the input (and

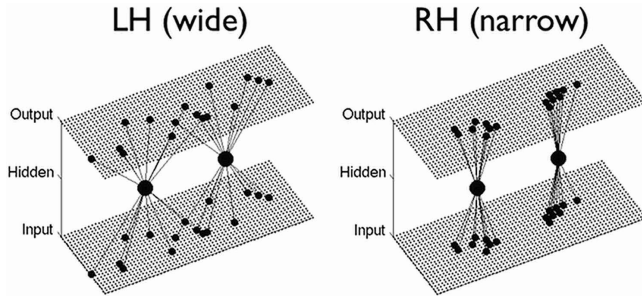


Fig. 1. Representation of two hidden units for LH (left) and RH (right) autoencoder networks, along with their connections. The connections are randomly sampled from a Gaussian distribution centered on each hidden unit's position in the input array. The Gaussian distribution used for the LH is wider than that used for the RH. Not pictured are the classification networks, which operate on the hidden unit encodings extracted from the autoencoder networks after training.

output) arrays are sampled from a Gaussian distribution centered at that hidden unit's position in the input (and output). The LH and RH autoencoders have the same number of hidden units and sample the same number of connections to the input (and output) for each hidden unit. The only difference between the networks, then, is the width (σ) of the Gaussian distribution.

In accordance with the findings of Galuske *et al.*⁴, the LH network has a wider distribution than the RH network (see Figure 1). Note that this differs from previous models of hemispheric differences, where the Gaussian describes the receptive fields of Radial Basis Function units¹⁵. In that work, wider receptive fields characterize the RH. Here, *whether or not there is a connection* is sampled from a Gaussian distribution. Our connections are sparse and there is no center-surround receptive field structure. Based on previous results in Hsiao *et al.*, we expected to find the opposite associations in this network architecture: a wider connection distribution would favor processing of HSFs, and a more narrow connection distribution would favor processing of LSFs.

The number of hidden units was set to the number of pixels (34×25 pixels = 850), to allow equal spacing across the input image with enough total parameters to learn the images. The number of connections per hidden unit (15) were fixed to be close to values reported in the literature for the number of interconnecting patches labeled by a single injection in V4^{1a}. Each LH and RH network

^aA wide range of values (smaller and larger) was tested for each of these parameters; results did not differ qualitatively when the number of connections per hidden unit varied to allow for the same number of overall weights to be used to learn the images. If too few weights were used, the networks could not learn the training set well enough for a meaningful analysis.

was constructed by randomly sampling connections for each hidden node. Gaussian distributions were used such that inter-patch distance values were similar to those reported in Galuske *et al.*, computed as the relative width of a “patch” (here, considered to be 1 pixel) vs. the relative distance between the center of patches (1.75 for the RH, 2.05 for the LH).

Table 1. The following four behavioral studies are addressed in the current modeling experiments.

Reference	Description
Sergent (1982)	Using Navon figures (large “global”-level letters composed of smaller “local”-level letters), this study showed a hemisphere \times level interaction, where target letters present at the larger, global level generally elicited a faster reaction times when presented to the LVF/RH, while target letters present at the smaller, local level generally elicited faster reaction times when presented to the RVF/LH.
Kitterle <i>et al.</i> (1992)	This study showed that the same stimuli used in different tasks could elicit different hemispheric advantages. This suggests that it is not the frequencies present in the stimuli alone that elicit the asymmetry in processing, but rather which frequencies are relevant to the task that drive the asymmetry. They found that a task requiring HSF information (square vs. sine waves, ignoring the frequency of the waves) elicited faster reaction times when stimuli were presented to the RVF/LH, while a task requiring LSF information (distinguishing two low frequency gratings, ignoring whether they are square or sine waves) elicited faster reaction times when stimuli were presented to the LVF/RH.
Christman <i>et al.</i> (1991)	This study showed that when two stimuli differed by a single spatial frequency, which hemisphere showed a processing advantage didn’t depend on the actual frequency of that component, but rather whether that frequency was relatively higher or lower compared to the other frequencies composing the stimuli. When the discriminative frequency was higher than the frequency content of the rest of the stimulus, responses were faster for presentation to the RVF/LH; when the discriminative frequency was lower than the frequency content of the rest of the stimulus, responses were faster for presentation to the LVF/RH.
Young & Bion (1981)	This study showed that accuracy on a face identification task was better when the probe faces were presented to the LVF/RH. It has been shown elsewhere that the RH may be dominant for face identification due to its advantage for low frequency processing ⁹ , so we choose to test our model on this task as well.

We aimed to model the four datasets described in Table 1. Each of these four datasets were collected using brief, lateralized presentation of an image, then requiring some classification of that stimulus (e.g. whether the stimulus contained

a target feature or not). We assume that asymmetries elicited in this paradigm are due to hemispheric differences in encoding these stimuli, and asymmetries in the classification of these stimuli are due to the asymmetries of the encodings. Our basic model setup, then, is to construct greyscale images for each task; to encode these images in our LH and RH autoencoder models with asymmetric connectivity patterns, and then to classify these asymmetric encodings in separate, but identical, LH and RH feed-forward neural networks trained on the classification task corresponding to the encoded images for that dataset.

For each dataset, greyscale images were constructed for each task stimulus. The autoencoders were trained via backpropagation of error¹⁶, using batch processing and the sum-squared error (SSE) cost function, to reproduce these greyscale images from the input to the output. Once the autoencoders reached a pre-determined performance level ($SSE = 0.02$, per output unit), training stopped. We used two types of regularization (weight decay (each iteration, every weight is multiplied by $\lambda = 0.975$) and adding noise (0.01%) to our training images) to encourage good generalization. Each stimulus image was presented to the trained autoencoder network, and the activation of the hidden units was recorded. These encodings, which differed only due to the differences in connectivity structure between LH and RH networks, were then used as inputs to a separate 3-layer feed-forward neural network, which was trained to classify these encoded stimuli according to the behavioral task given to human subjects in the original experiment. The number of hidden units was different for each classification network (but identical between hemispheres), and was chosen such that the classification task could be learned well. We chose to separate general visual encoding (our autoencoder network) from task performance (our perceptron), to reflect the idea that retinotopic visual encoding is established during early visual experience and does not change during task performance. In subsequent work, we train our autoencoder networks with natural images to further substantiate this idea.

For a single experiment, multiple instances of each LH and RH network were constructed and trained; each instance differs only in the random sampling of its connections.^b Performance was evaluated on each model individually, then performance for all instances of each hemisphere were averaged. Average model error for each model hemisphere was compared to average reaction time in the corresponding human experiment, with both conceived as measures of difficulty or uncertainty in processing.

^bThe number of model instances was chosen such that the total number of trials across all stimuli and model instances matched that reported in the corresponding behavioral experiment, to match statistical power of results.

In order to examine how the different connectivity distributions affect spatial frequency encoding, each stimulus image was presented to a trained autoencoder. Each output image produced was examined for spatial frequency content, and a 2D spectrogram across all images in the stimulus set was constructed. Each 2D spectrogram was translated to a 1D spectrogram.^c Each spectrogram was compared to the spectrogram of the original image; the absolute value of the difference between the spectrograms represents the error for each model hemisphere’s encoding at each spatial frequency. This error measure was then compared across hemispheres, showing for each frequency which hemisphere has encoded information closer to the original image than the other.

2.2. Results

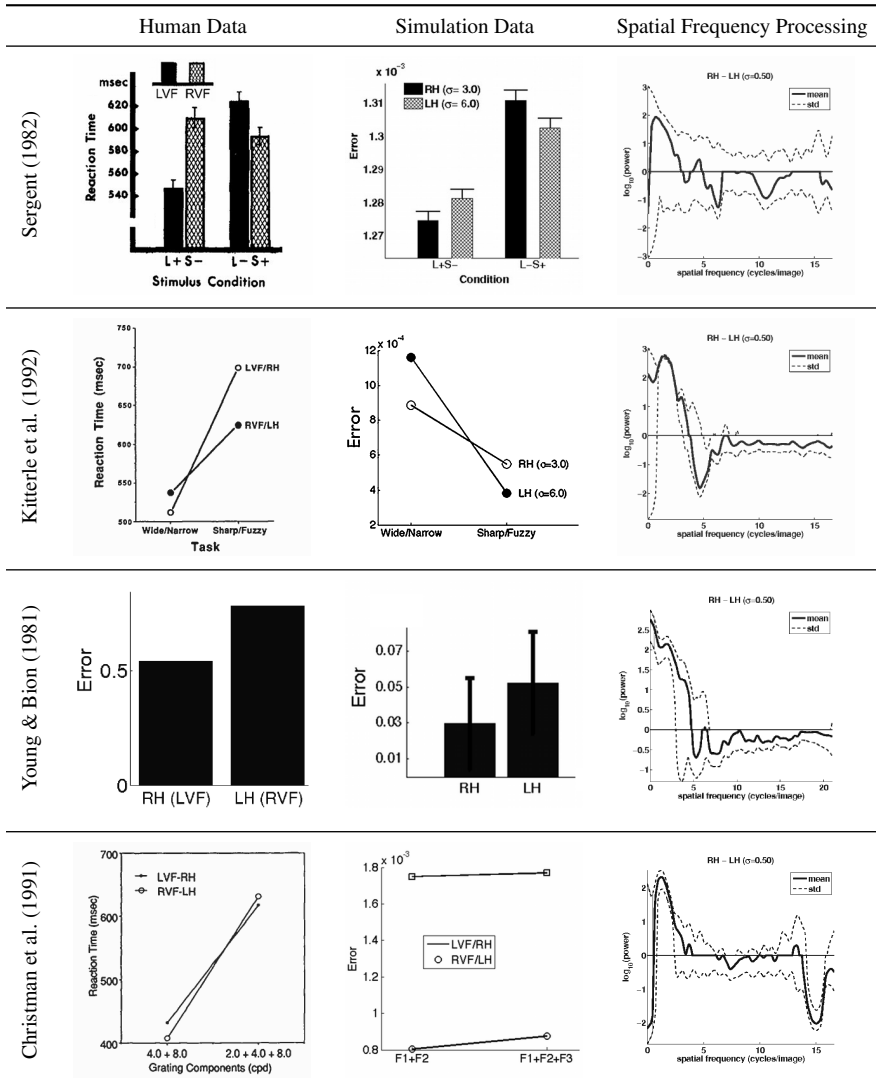
Results are summarized in Table 2, showing that for three of the four tasks, the model performance (column 2) showed the interaction or performance differences found in the human behavioral measures (column 1). In addition, in all four tasks, the model shows an interaction between frequency band and hemisphere (column 3): greater accuracy in reproducing images for the RH networks (above zero-line) is focused in the lower spatial frequencies (towards origin), while greater accuracy for the LH networks (below zero-line) is focused in the higher spatial frequencies (away from origin).

This indicates that, somewhat counterintuitively, narrowly distributed connections are biased relatively more towards LSF processing, and widely distributed connections are biased relatively more towards HSF processing. We are investigating the cause of this relationship in current work. Preliminary results indicate that average spacing between nearest-neighbors of input/output units connecting to a single hidden unit governs this behavior. Specifically, when connection spread is small, constructive/destructive interference patterns between LSF vs HSF gratings of different phases and orientations occur: LSF gratings shows large variations in hidden unit activation based on particular orientations and phases, while HSF gratings do not. This seems to allow for more specific encoding of shapes and contours in LSF gratings, and thus a bias towards representing LSF information.

Examining these results more closely, we found two places where our models did not accurately reflect the human behavior. In our replication of Kitterle *et al.*, while we did show the measured hemisphere \times task interaction, the relative difficulty of the two tasks did not match that in the human measures. This diffi-

^cEach 2D frequency (f_x, f_y) was converted to a 1D frequency $(f_r = \sqrt{f_x^2 + f_y^2})$; for each unique value of f_r , all power/amplitude values were then averaged to create the 1D spectrogram.

Table 2. Simulation results, comparing human & simulation output



culty may depend on different sensory transformations, or other issues. In fact, in subsequent work, we have found this to be true in the modeling; changing the training dataset, or regularization parameters, does not affect the hemisphere × task interaction, but does affect the relative difficulty of the two tasks. Therefore, we don't think that this mismatch is a critical issue to consider further here.

We believe the second issue we found is more important, and more interesting. Our models did *not* show a hemisphere \times stimulus class interaction for the Christman *et al.* study. We tested a few possible explanations for this. We varied the three base frequencies that composed the stimuli; this affected which model hemisphere showed better performance, but no reliable hemisphere \times stimulus class interaction was found. We tried larger images, to expand the range of spatial frequencies that could be encoded and obtained similar results. Lastly, we trained the autoencoder on a separate dataset, then extracting hidden unit encodings on the task-relevant stimuli. Again, this did not show any consistent interaction.

Using data from our developmental model, we discuss an alternative interpretation of the Christman *et al.* study in the Discussion section below.

3. Patchy connection networks may develop asymmetrically in humans, under normal visual development

The results above clearly suggest an association between narrowly distributed connections and LSF processing, and widely distributed connections and HSF processing. Here, we investigated whether a developmentally-inspired model can learn these connection patterns through experience. The model incorporates three observations/constraints: 1) The right hemisphere begins maturing earlier than the left; 2) early in development, the input to the system is concentrated in the low spatial frequencies, due to low acuity in infancy; and 3) synaptic connections are subject to a pruning process to remove weak or ineffective connections.

This setup has strong parallels with human visual development. Hellige⁷ has provided evidence that the RH begins maturing earlier than the LH. We know that visual acuity changes rapidly in infants, such that visual acuity and sensitivity to high frequencies increases over time. This means that the RH will mature while visual input is more biased towards low frequencies, and the LH will mature while the visual input is less biased. Katz and Callaway¹² have shown that connections between patches mature by strengthening connections consistent with visual input, and pruning (removing) connections that are not. Thus, our goal is not only to try and corroborate our previous findings, but to provide a plausible neurodevelopmental account of where both anatomical and behavioral asymmetries come from.

3.1. Methods

We created two sets of autoencoder neural networks, each completely identical. They each had the same connectivity pattern, consisting of 20 connections selected from a circular 2D Gaussian distribution from the hidden units to the input

and output units, with a fixed number of connections from each hidden unit to the input and output layers. The only difference between these sets of neural networks was that one was trained on blurred, low-frequency-passed images (our RH networks) and the other was trained on full-fidelity images (our LH networks), to simulate differences in development. As a first pass, we used a Gaussian kernel ($\sigma=8\text{px}$) to blur our images for the RH network.

Networks were trained using back propagation of error (batch mode, using SSE as described above), to simulate the strengthening of connections through learning. To simulate the pruning aspect of development, after every 10 epochs, we removed a constant percentage of the weakest connections (13.4%), such that after 50 epochs exactly 50% of the connections remained. At the end of training, each network had a 50% reduction in its original connections, such that on average, each hidden unit had 10 connections.

3.2. Results

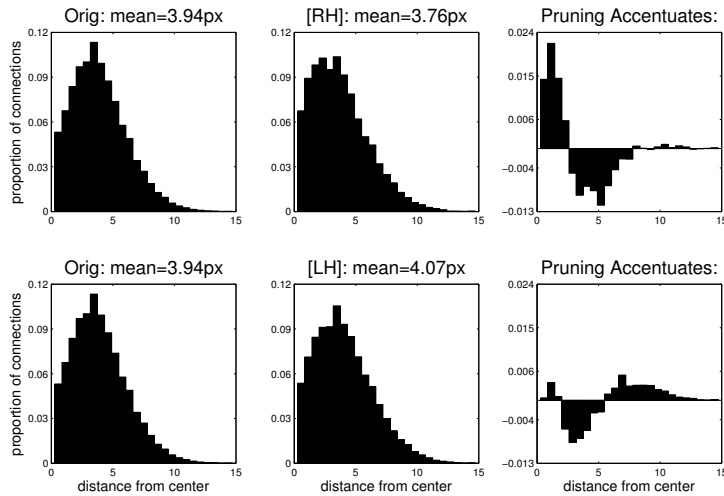


Fig. 2. Distribution of connections for networks trained on blurry/low-pass (RH; top row) and full-fidelity (LH; bottom row) images, as well as the difference between each of those distributions from the original distribution.

After training, the network trained on low-pass/blurry images (RH) biased the original distribution towards shorter connections, while the network trained

on full-fidelity images (LH) biased the original distribution towards more distant connections (see Fig. 2). This corroborated the associations we saw in our first study, where a smaller spread of connections favored storing information in LSFs, and where a wider spread of connections favored storing information in HSFs. Analysis of the spectral power differences between the networks also showed a similar bias as described in Table 2.

We are examining the computational reasons behind this association in current work.

4. Reprising the “relative frequency” effect

Christman *et al.*³ showed that when a frequency grating alone distinguished two stimulus classes, the effect of that grating depended on the frequency of that grating *relative* to the other frequencies that made up those stimuli. When the distinguishing grating was higher than the rest of the frequencies used in the stimuli, the stimulus containing the distinguishing grating showed a LH advantage; when the distinguishing grating was lower, there was a RH advantage. Ivry and Robertson propose that this “relative frequency” effect is itself a computational mechanism: filters at different spatial scales / frequency bands are weighted and summed based on the task; a top-down bias to those weights (that differs between the hemispheres) is relative to the the frequencies selected by those weights.

An alternate hypothesis is that the same filters that select the task-based frequency range are themselves different between the hemispheres. Then, when those filters are weighted and summed based on task demands, a bias that is dependent on that weighting occurs. This alternate hypothesis is not possible in the DFF, where the filters represent bottom-up processing in all conditions (low and high stimulus strength), and therefore would predict hemispheric differences in every experiment. This alternate hypothesis is possible with our model, however. We argued that long-range lateral connections are only active during low stimulus strength, as found in the experiments detailed above. In our case, different “filters” would be different cortical areas, each of which process information at different spatial scales.

Looking to the literature, and examining unpublished simulations, we find data consistent with our proposal. Tasks at different spatial scales do engage different cortical areas⁸, consistent with our suggestion that cortical areas can act as task-based filters of spatial scale / frequency. Using this idea in our developmental model suggests that homologous cortical areas between LH and RH would have different frequency biases *relative to the spatial scale of that area*. In that model, we found that small reductions (due to synaptic pruning) in the average spread of

long-range lateral connections biases processing towards LSFs. Further experiments with our developmental model showed that, the greater the original spread of connections, the lower the frequency at which the RH stopped showing an advantage (the cross-over point, where RH and LH performance is similar). Thus, in our proposal, homologous areas with larger spatial scales would have to have a larger average spread of connections in order show a frequency bias relative to that spatial scale. In fact, that's exactly what's found in the macaque monkey: areas higher in the visual processing hierarchy show longer average connections, both in cortical distance, and in the percentage of the visual field they span over¹. Our modeling data also suggest that, given the small average spread found in early visual areas, little or no frequency biases could develop there. This is consistent with the idea that visual processing asymmetries are found beyond early sensory processing.

Thus, the failure of our model to account for Christman *et al.*'s data may indicate that these data represent processing in two different cortical areas, each selective for information at a particular spatial scale, each with frequency processing biases relative to that spatial scale. Our autoencoder model above represents a single cortical area; to model this, we would need multiple autoencoder models, each with information at different spatial scales (possibly different image resolutions) and with corresponding changes to the average connections spread (greater for images of lower resolutions).

5. Discussion

Here we showed that an asymmetry in a specific type of intra-cortical connectivity can account for local/global behavioral data, face processing data, and matches spatial frequency asymmetries reported in the literature. This model provides a biologically grounded implementation for these phenomena, and the analyses here showing consistent frequency filtering differences in the model hemispheres are consistent with the current algorithmic explanation for visual processing asymmetries via frequency filtering. These frequency filtering differences are found at a post-sensory encoding stage, consistent with the suggestions in the literature.

We have also implemented a plausible developmental scenario for the emergence of this connectivity asymmetry, which uses three principles supported by developmental data (changes in visual acuity over development, both strengthening and pruning of long-range lateral connections during development, and different time-courses for RH and LH maturation).

Finally, we have suggested that our failure to model the results of Christman *et al.* with a model of a single cortical area is not a failure of our model, but instead

suggests that the “relative frequency” effect in visual processing asymmetries is itself not a mechanism in visual processing. We propose that this effect is the result of latent biases within higher areas in the visual hierarchy that are established developmentally and are relative to that area’s spatial scale of processing. The bias of an area affects visual processing when the spatial scale of processing for a task matches that of the area, and when stimulus strength is sufficiently weak to activate these intrinsic long-range lateral connections.

Acknowledgments

This research was funded in part by a CARTA (Center for Research and Training in Anthropogeny) Fellowship, and by NSF grant SMA 1041755 to the Temporal Dynamics of Learning Center.

References

1. Amir, Y., Harel, M. and Malach, R. (1993). Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex, *The Journal of Comparative Neurology* **334**, 1, pp. 19–46.
2. Christman, S. (1989). Perceptual characteristics in visual laterality research, *Brain and Cognition* **11**, 2, pp. 238–257.
3. Christman, S., Kitterle, F. L. and Hellige, J. (1991). Hemispheric asymmetry in the processing of absolute versus relative spatial frequency, *Brain and Cognition* **16**, 1, pp. 62–73.
4. Galuske, R. A., Schlote, W., Bratzke, H. and Singer, W. (2000). Interhemispheric asymmetries of the modular structure in human temporal cortex, *Science (New York, N.Y.)* **289**, 5486, pp. 1946–1949.
5. Gilbert, C. D. (1992). Horizontal integration and cortical dynamics, *Neuron* **9**, 1, pp. 1–13.
6. Grossberg, S. and Williamson, J. R. (2001). A neural model of how horizontal and interlaminar connections of visual cortex develop into adult circuits that carry out perceptual grouping and learning, *Cerebral Cortex* **11**, 1, pp. 37–58.
7. Hellige, J. B. (1993). *Hemispheric Asymmetry: What’s Right and What’s Left* (Harvard University Press), ISBN 0674005597.
8. Hopf, J., Luck, S. J., Boelmans, K., Schoenfeld, M. A., Boehler, C. N., Rieger, J. and Heinze, H. (2006). The neural site of attention matches the spatial scale of perception, *The Journal of Neuroscience* **26**, 13, pp. 3532–3540.

9. Hsiao, J., Shahbazi, R. and Cottrell, G. (2008). Hemispheric asymmetry in visual perception arises from differential encoding beyond the sensory level, *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*.
10. Hsiao, J. H.-w., Shieh, D. X. and Cottrell, G. W. (2008). Convergence of the visual field split: Hemispheric modeling of face and object recognition, *Journal of Cognitive Neuroscience* **20**, 12, pp. 2298–2307.
11. Ivry, R. B. and Robertson, L. C. (1998). *The Two Sides of Perception* (The MIT Press), ISBN 0262090341.
12. Katz, L. C. and Callaway, E. M. (1992). Development of local circuits in mammalian visual cortex, *Annual Review of Neuroscience* **15**, 1, pp. 31–56.
13. Kitterle, F. L., Hellige, J. B. and Christman, S. (1992). Visual hemispheric asymmetries depend on which spatial frequencies are task relevant, *Brain and Cognition* **20**, 2, pp. 308–314.
14. Levitt, J. and Lund, J. (2002). Intrinsic connections in mammalian cerebral cortex, in A. Schuez and R. Miller (eds.), *Cortical areas: unity and diversity* (CRC Press), ISBN 9780415277235, pp. 133–154.
15. Monaghan, P. and Shillcock, R. (2004). Hemispheric asymmetries in cognitive modeling: Connectionist modeling of unilateral visual neglect, *Psychological Review* **111**, 2, pp. 283–308.
16. Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986). Learning representations by back-propagating errors, *Nature* **323**, 6088, pp. 533–536.
17. Sergent, J. (1982). The cerebral balance of power: confrontation or cooperation? *Journal of Experimental Psychology. Human Perception and Performance* **8**, 2, pp. 253–72.
18. Sergent, J. (1983). Role of the input in visual hemispheric asymmetries, *Psychological Bulletin* **93**, 3, pp. 481–512.
19. Sergent, J. (1985). Influence of task and input factors on hemispheric involvement in face processing, *Journal of Experimental Psychology. Human Perception and Performance* **11**, 6, pp. 846–861.
20. Swadlow, H. A. and Alonso, J. (2009). Spikes are making waves in the visual cortex, *Nat Neurosci* **12**, 1, pp. 10–11.
21. Young, A. W. and Bion, P. J. (1981). Accuracy of naming laterally presented known faces by children and adults, *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* **17**, 1, pp. 97–106.