

Categorical vs Coordinate Relationships do not reduce to spatial frequency differences

Vishaal Prasad (v2prasad@ucsd.edu)

UCSD Department of Computer Science, 9500 Gilman Dr.
La Jolla, CA 92093 USA

Ben Cipollini (ben.cipollini@gmail.com)

Classy.org, Diamond View Tower, 350 Tenth Ave
San Diego, CA 92101 USA

Garrison W. Cottrell (gary@ucsd.edu)

UCSD Department of Computer Science, 9500 Gilman Dr.
La Jolla, CA 92093 USA

Abstract

Categorical and coordinate stimulus processing were hypothesized by Kosslyn (1987) to be lateralized visual tasks, differentiated by task-relevant spatial frequencies. Slotnick et al. (2001) directly tested Kosslyn's hypothesis and concluded that the lateralization presents only when tasks are sufficiently difficult. Our differential encoding model is a three layer neural network that accounts for lateralization in visual processing via the biologically plausible mechanism of differences in connection spread of long-range lateral neural connections (Hsiao, Cipollini, & Cottrell, 2013). We show that our model accounts for Slotnick's data and that Slotnick's analysis does not convincingly explain their results. Instead, we propose that Kosslyn's initial hypothesis was based on an incorrect assumption: categorical and coordinate stimuli are not solely differentiated by spatial frequencies. The results that our model captures cannot be reproduced by Ivry and Robertson's (1998) Double Filtering by Frequency theory, which is driven solely by lateralized spatial frequency processing.

Keywords: Differential encoding; hemispheric asymmetry; spatial frequency processing; categorical vs. coordinate

Introduction

The human brain is composed of two largely disconnected hemispheres that communicate via a bridge of neural connections known as the corpus callosum. The level of relative disconnection and redundancy suggests that, for some cognitive processes, it may be advantageous for each hemisphere to specialize and reduce redundancy. This functional specialization, or "lateralization," occurs in many diverse cognitive facilities in humans (Stephan et al., 2003) and non-humans alike (Rogers & Andrew, 2002). Examples in humans include fine motor skills and language processing, both of which are left hemisphere dominant (Knecht et al., 2000). Of particular interest to us is visual lateralization. Past studies have shown visual lateralization in processing stimuli ranging from frequency gratings to facial recognition (e.g. Ivry & Robertson, 1998; Sergent, 1985).

Navon's (1977) hierarchical letters are one visual stimulus showing lateralization. These consist of a large, "global" letter (e.g., "T") that is composed of small, "local" letters (e.g., "F"). Sergent (1982) showed an advantage in identifying the local level target (the letter F in the above example)

presented when flashed in the right visual field/ left hemisphere (RVF-LH), and the reverse (the letter T in the above example) in the left visual field/ right hemisphere (LVF-RH). She concluded that the LH performs better with high spatial frequency (HSF) stimuli, whereas the RH does better with low spatial frequency (LSF) stimuli. Kitterle, Christman, and Hellige (1990) directly tested this hypothesis with frequency gratings and showed that the lateralization was driven by task demands, rather than purely by stimulus properties.

Inspired by Sergent's (1982) theory, Ivry and Robertson (1998) proposed their Double Filtering by Frequency (DFF) theory to explain these asymmetric processing results. DFF theory proposes that the hemispheres identically first select the frequency bands relevant to the task, but then are biased so that the left hemisphere preferentially processes HSFs, and the right hemisphere LSFs. DFF theory also accounts for data suggesting that frequency processing differences between the hemispheres are not absolute, but instead are relative to the frequency band relevant for solving the task (Christman, Kitterle, and Hellige, 1991). Finally, their model accounts for the categorical and coordinate spatial relations proposed by Kosslyn (1987) and Kosslyn, Koenig, Barrett, Tang, and Gabrieli (1989), described below. However, there is no neurological basis for the core mechanisms of the DFF theory, nor is there a developmental explanation of how or why this phenomenon would emerge (see Cipollini, 2014 for further discussion).

Kosslyn (1987) and Kosslyn et al. (1989) argued that humans process visual stimuli using two distinct types of spatial relations. Coordinate relations rely on an absolute, metric basis; for example, the statement "the glass of water is 3 inches from my hand" defines a coordinate judgment of one's hand and the glass of water. In contrast, categorical relations rely on abstract, relative terms. The statement "the glass of water is on top of the table" does not tell us exactly where the glass is, only its relative position to a table. In his 1987 paper, Kosslyn observed a RVF-LH advantage for categorical relation judgments and a LVF-RH advantage for coordinate relation judgments in response time. Other work (e.g. Hellige & Michimata, 1989) provided further support for this hy-

pothesis with more varied types of stimuli (e.g., a bar and dot stimulus).

However, Sergent (1991) found that this lateralization effect presented only when stimuli were degraded, and several analyses have noted that the lateralization only presents in right-handed people (Slotnick, Moo, Tesoro, & Hart, 2001). Other researchers likewise found weak or inconclusive evidence for lateralization of categorical and coordinate stimuli, especially for the LH advantage on categorical stimuli (Okubo & Michimata, 2002). Nevertheless, researchers generally agree that a distinction exists, even if it is weaker than originally thought (see Jager & Postma, 2003 for a review).

In light of these conflicting data, Kosslyn and colleagues now argue that lateralization in categorical and coordinate stimuli results from a differential frequency processing, potentially based on lateralization in neuronal receptive fields (see Baker, Chabris, & Kosslyn, 1998; Kosslyn, Chabris, Marsolek, & Koenig, 1992). In support of this hypothesis, Okubo and Michimata (2002) showed that the RH coordinate advantage, but not the LH categorical advantage, was eliminated by contrast balancing, which removes low spatial frequency information without degrading stimuli.

Slotnick et al. (2001) revisit Categorical / Coordinate

Of particular interest here are the experiments and results of Slotnick et al. (2001). In past experiments, researchers flashed stimuli in one visual field or the other, leaving room for interhemispheric interference. To isolate lateralization effects, Slotnick et al. ran a clinical study on 134 subjects, each of whom had at least one hemisphere temporarily deactivated by an intracarotid injection of sodium amobarbital.

Using the same stimuli as in Kosslyn et al. (1989), Slotnick et al. (2001) sought to reproduce their results in subjects with deactivated hemispheres. In addition, they added a new stimulus type, paired squares, which was designed to resist “categorization” of coordinate tasks, whereby a subject on later trials during an experiment learns a coordinate task (e.g., is the plus more than two inches from the minus) and turns it into a categorical task (Slotnick et al., 2001). This explanation had been proposed to explain the weakening of the RH advantage on coordinate stimuli. The paired squares coordinate stimulus forces the subject to make a direct metric comparison between the two parts of the stimulus.

Unlike the original paper, Slotnick et al. (2001) measured results by classification error, rather than reaction time. Though the results generally aligned with Kosslyn’s hypothesis, one coordinate experiment did not show the expected RH dominance, instead showing lateralization opposite of that in the original paper. The authors noted that distances between components of their figures (e.g. the blob and dot), were larger in their experiments than in Kosslyn et al. (1989). Difficulty has been reported to modulate lateralization in other experiments (e.g. Sergent, 1985), and so they posited this made the task too easy to show proper lateralization. Consequently, post-hoc they stratified the tasks by difficulty and found that only when a task is sufficiently difficult does later-

alization arise as expected.

Differential Encoding Model

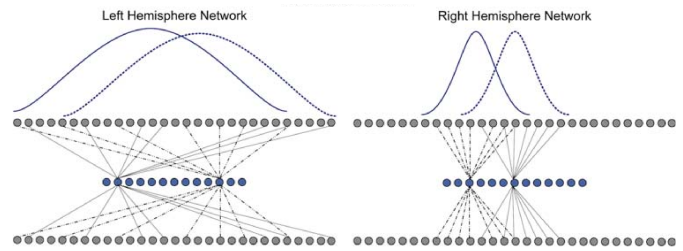


Figure 1: Taken from Hsiao, Cipollini, and Cottrell (2013), this diagram shows the autoencoder models with varying connection spreads and symmetric connections. Notice the left hemisphere’s hidden units connect to a more spread out set of neurons on average, while maintaining the same number of connections.

Competing with DFF theory is our Differential Encoding (DE) theory (Hsiao, Shahbazi, & Cottrell, 2008; Hsiao, Cipollini, & Cottrell, 2013). It is inspired by an anatomical difference in the auditory system’s long range lateral connections (LRLCs). On average, a LH neuron connects to neighbors generally farther from itself than the RH neurons do (Galuske, Schlote, Bratzke, & Singer, 2000). The DE model hypothesizes these LRLCs as the driving factor behind visual lateralization as well. Compared to the DFF theory, the DE model has the advantage of having neurodevelopmental and neuroanatomical plausibility (Cipollini, 2014).

Computationally, the Differential Encoding model is a standard 3-layer neural network which can be thought of as a recurrent neural network unrolled one step in time. The first set of connections is a sparse autoencoder, trained on natural images, to represent how a stimulus might be transformed in the early stages of the brain using low level processing such as Gabor filters. Each neuron in the autoencoder corresponds to a spatial location, and it connects to 5 other neurons generated randomly from a Gaussian centered around the neuron itself. The LH and RH networks vary by the standard deviation, or sigma parameter, of the Gaussian, to mimic the lateralized connection spread of the LRLCs, seen in Figure 1.

Note that this differs from a Gaussian receptive field. A network with Gaussian receptive fields has fixed connections, and the strength of the connections are determined by a Gaussian. In the DE model, the lateral connections themselves are stochastic and determined by randomly sampling from a Gaussian, and the strength of the connections are learned by training.

In Hsiao, Cipollini, and Cottrell (2013), the authors show that the autoencoder in the DE model reconstructs natural images in accordance with the predictions of Sergent (1982). Specifically, the RH model reconstructs low spatial frequency (LSF) components of a stimulus better, whereas the LH model reconstructs HSF components better.

Once trained, each hemisphere’s hidden units are then con-

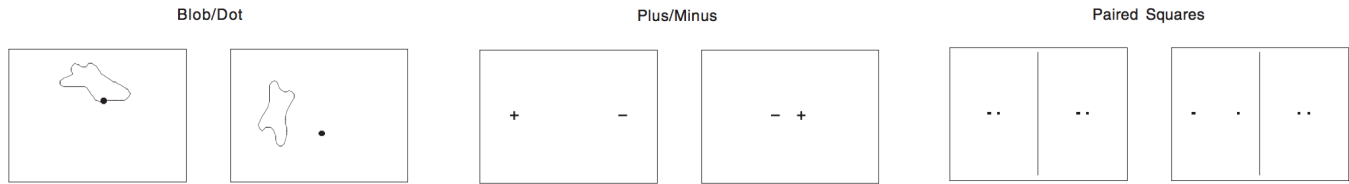


Figure 2: These are the stimuli from Slotnick et al. (2001). Note that paired squares only had a coordinate task, whereas blob/dot and plus/minus have both categorical and coordinate.

nected to a task-specific output unit that is trained by the delta rule to learn some task. In this way, the information represented by the hidden layer is tested as to what tasks it is best at. We have found in many experiments that the LH model is better at tasks that require HSFs, and vice-versa for the RH model (Hsiao, Cipollini, & Cottrell, 2013).

In addition to the autoencoder properties outlined above, the model has accounted for Sergent’s (1982) data, as well as Kitterle, Christman, and Hellige’s (1990) data showing task dependence of lateralization (Hsiao, Cipollini, & Cottrell, 2013). This suggests the DE model has the very frequency encoding properties that Sergent (1982) hypothesized. As Kosslyn and colleagues have suggested (e.g., Baker, Chabris, & Kosslyn, 1999), the distinction between categorical and coordinate stimuli may stem from lateralized frequency processing. Therefore, we test the network on Slotnick et al. (2001)’s stimuli to further establish the relationship between our model and frequency lateralization, as well attempt to reach parity with the DFF on these stimuli.

Materials and Methods

Our stimuli mimic those of Slotnick et al. (2001)

The stimuli used in the 2001 study can be seen in Figure 2. There are three types of stimuli: blob/dot, plus/minus, and paired squares. All three stimulus types involve coordinate tasks. Blob/dot requires an evaluation of how far apart the blob and dot are, and plus/minus likewise requires an evaluation of how far apart the plus and minus are. The paired squares task, in contrast, requires judging whether the two sets of paired squares are equidistant or not. The former two stimulus types also have categorical tasks. The blob/dot categorical task requires evaluating whether the dot is on the blob or off of it, and the plus/minus categorical task requires evaluating whether the plus is on the right or the left. There is no categorical task for paired squares.

For the plus/minus and blob/dot coordinate stimuli, “near” configurations were those where the distance between the plus and minus or blob and dot measured smaller than a reference distance of 2 inches; the “far” configurations were larger than 2 inches. In our model, the reference distance was 5.5 and 6 pixels for plus/minus and blob/dot respectively.

Slotnick et al. (2001) hypothesized that these tasks are harder when the distance between stimulus components are close to the reference distance of 2 inches. They defined “hard” configurations as those where the distance between

stimulus components fell within the range of [1, 3] in inches, i.e. within one inch from the reference distance of 2 inches. “Easy” configurations were those outside of this range. In our model, “hard” configurations were those where the distance fell within 2 pixels of the reference distance, and “easy” were the other stimuli. Stratifications for paired squares coordinate and the two categorical tasks were less principled and will be explored below.

Simulation Procedure

The simulation was implemented in MATLAB. All code is open source¹. Input images were implemented as bitmaps, following the images published in the original paper as best possible. To accomplish this, plus/minus and paired squares images were 34x25 pixels. Due to the need for increased resolution, blob/dot images were 68x50 pixels. Due to the varying resolutions, the experiment sizes had different hyperparameters. Of particular interest, the 34x25 images had a RH and LH standard deviation (sigma) of 4 and 10 pixels respectively; the 68x50 had 4 and 15. In both cases, each neuron had five connections per hidden unit, with one hidden unit corresponding to each pixel of the image. Our train and test data were the same, so to prevent overfitting, we used heavy regularization. Specifically, we used a relatively high amount of dropout of 0.7 (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014) and introduced noise on the input.

The human experiment used 100 LH subjects and 124 RH subjects; however, 54 hemispheres of patients deemed abnormal or otherwise compromised (e.g. those with parietal lobe tumors) were excluded (Slotnick et al., 2001). We followed the same analyses done in the human experiment, and in an attempt to roughly match statistical power, we instantiated each hemisphere in our computational model 100 times. Instantiations with outlier performance were discarded, so there were slightly fewer than 100 instantiations for the final results.

Results and Discussion

The output of the DE model is a real-valued number between 0 and 1, where 0 and 1 represent the target labels (e.g. off/on for categorical blob/dot). Error is measured as the sum-squared error (SSE) between the model’s output and the true label. In Slotnick et al. (2001), they measure mean percent error in classification. The different measurements means y-

¹<https://github.com/guruucsd/DifferentialEncoding/releases/tag/slotnick>

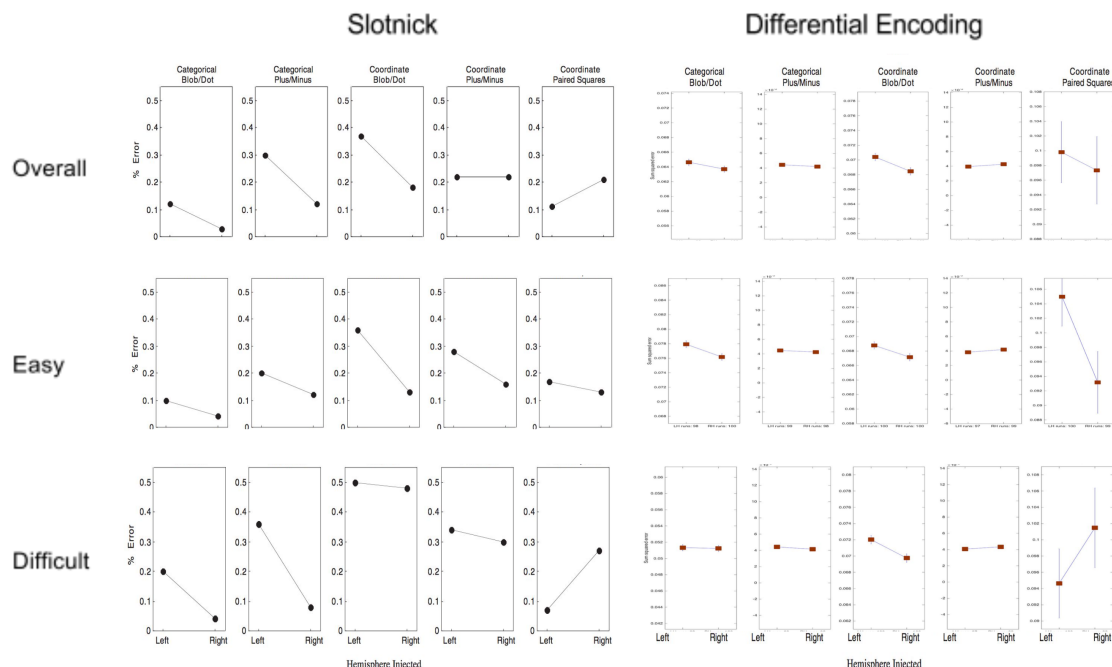


Figure 3: Differential Encoding (left) results follow the overall Slotnick et al. (2001) results for task x hemisphere interactions, which can be seen as the slope of each line. Note that since hemispheric performance, not absolute performance, was relevant, y-axes were re-scaled to emphasize slope.

axes are not directly comparable, but in this experiment we are concerned with the relative performance of hemispheres on each task. We simply compare the slopes in Figure 3 to see how well the DE model fits the human data.

From Figure 3, it is clear that the DE results do not follow the human data perfectly, but the key concepts are captured. Categorical blob/dot, coordinate blob/dot, and categorical plus/minus all showed the expected LH advantage. Coordinate plus/minus shows no LH advantage ($F_{1,195}=0.0669$, $p > 0.75$), and categorical plus/minus has a stronger LH lateralization than it ($F_{1,391}=12.96$, $p < 0.001$). Crucially, the anomalous result in the original paper persists: categorical blob/dot is not more LH-dominant than coordinate blob/dot ($F_{1,396}=1.200$, $p > 0.25$), as it was in Kosslyn et al. (1989).

Paired squares on first glance appears to have reverse lateralization as in the human data. However, paired squares was extremely volatile, and that advantage disappears or reverses spontaneously. Statistically, lateralization was non-significant: there are large error bars for paired squares, and the repeated-measures anova of this data reveals this same non-significance ($F_{1,198}=0.115$, $p > 0.9$). Slotnick et al. (2001) reported this stimulus had only marginal statistical significance ($p < 0.1$) in their results as well. Therefore, while we plan to investigate this stimulus further, for now, we are less concerned that it did not show LH dominance.

The middle and bottom rows in Figure 3 show the results of the human data and the DE model for easy and hard subsets respectively. Our model replicates the results well. Before further analyzing the results, we take a closer look at Slotnick

et al.'s (2001) difficulty stratification, to understand what relationships across difficulty are crucial to replicate.

Revisiting Slotnick et al. (2001)

Slotnick et al. (2001) directly measured lateralization in subjects who had a hemisphere temporarily deactivated as part of a presurgical evaluation for treatment of intractable epilepsy. This meant the authors only ran their experiments once. When their data contradicted Kosslyn's results, they ran a post-hoc analysis of the data to explain their results. The crucial takeaway was that lateralization presented only if the task is difficult enough.

We find reasons to doubt Slotnick et al.'s conclusions. First, there are critical inconsistencies in their figures. If the easy and hard instances of a task both lateralize in the same direction, then combining all trials together should as well. Yet, as shown in Figure 3, coordinate plus/minus overall does not lateralize, though its easy and difficult subsets did. This discrepancy is never addressed in their paper.

In addition, most of the stratifications between easy and hard were not built in a principled manner and therefore lack validity. Slotnick et al. (2001) state that the stratification of the paired squares task was an empirical heuristic, as there was no neat way of differentiating easy and hard stimuli. Similarly, they note that no analogous concept of difficulty exists for categorical stimuli, so they just used the same division as their coordinate counterparts.

The coordinate blob/dot results are both internally consistent and well-principled, but the other tasks are not. The overall results of Slotnick et al. (2001), matched by our compu-

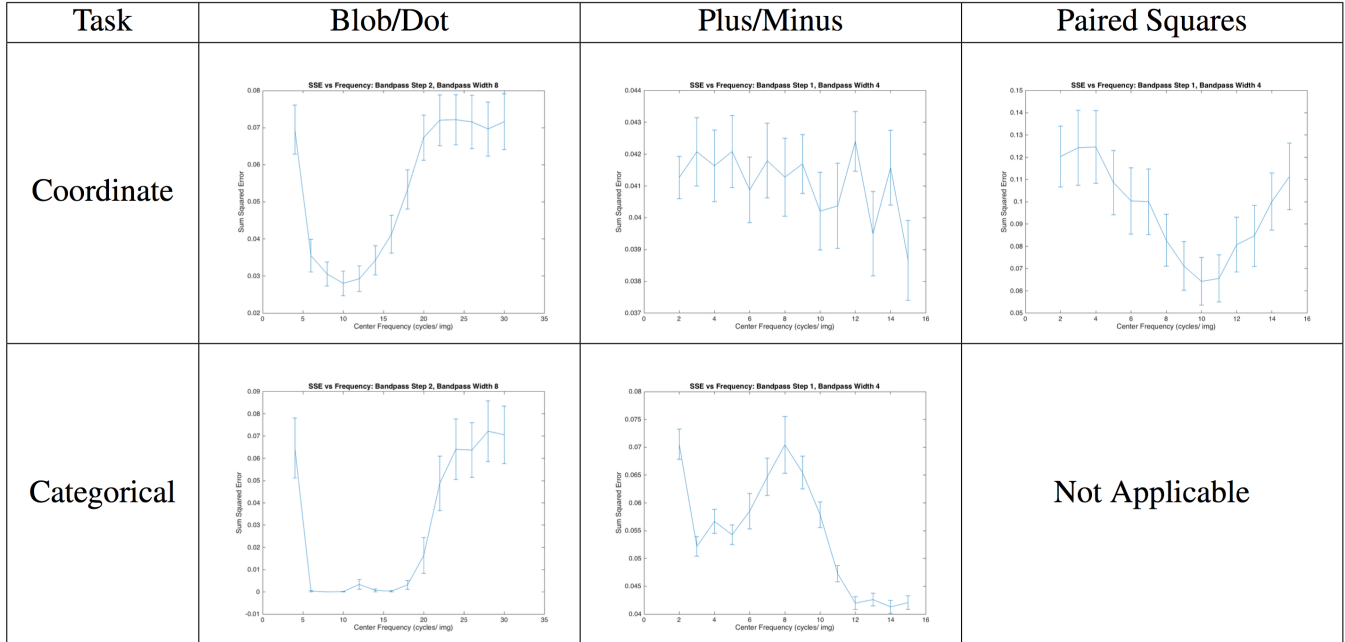


Figure 4: Frequency preferences of the DE model for each task. Y-axis is network accuracy, while x-axis marks the center of the frequency window on which a bandpass filter is applied. If spatial frequencies drive categorical and coordinate processing, we should see better performance on categorical tasks at HSFs and vice-versa at LSFs. We do not see this pattern.

tational modeling, are not adequately explained by difficulty stratification. We now look for an alternate explanation.

Spatial frequency selectivity within the DE model

Kosslyn and colleagues (e.g. Baker, Chabris, & Kosslyn, 1999) concluded that categorical and coordinate processing lateralized according to preferential frequency processing. As originally hypothesized in Sergent (1982), the LH is thought to outperform the RH in processing HSFs and vice-versa for LSFs. Numerous experiments have shown lateralization as a function of filtering stimuli to specific frequency windows (e.g. Sergent, 1985). The Differential Encoding model has also shown this same differential frequency processing (Hsiao, Cipollini, & Cottrell, 2013).

To examine whether spatial frequency differences drove results on these five tasks, we tested the model with different bandpass filters for each task. Specifically, all networks trained on the same, unchanged image patches to learn the same features, simulating typical visual experience. However, the perceptron was trained and tested on stimuli run through a bandpass filter window of size four and eight CPI, for 34x25 and 68x50 images respectively. This would allow us to empirically deduce which frequencies the network best responded to. Results were agnostic to a host of parameter choices, including sigma, dropout, and bandpass width (within reason), so we believe the results are general to the task, and not specific to anything about our network setup.

Per Kosslyn’s hypothesis, we expected to see lateralization in accordance with task type: there should be increased categorical performance on HSFs and likewise for coordinate and LSFs. Figure 4 shows the results. Coordinate paired squares

is almost parabolic with a minimum around 10 CPI. Coordinate plus/ minus is largely agnostic to frequencies, whereas categorical shows bimodal preference, with the global minimum at HSFs. Categorical blob/dot performed equally well at the windows centered from 10 to 18 CPIs, whereas coordinate performed best in the window centered at 10 CPIs and was locally parabolic around that area. Beyond roughly 20 CPI, critical image features are lost and in both cases the networks perform poorly.

There is no unified trend of stimulus type and frequency. Considering the DE model both captures the critical relationships outlined in Slotnick et al. (2001) and accounts for spatial frequency filtering (Hsiao, Cipollini, & Cottrell, 2013), it appears something besides frequency underlies the categorical and coordinate dichotomy. We plan to explore this further via contrast balancing (Okubo & Michimata, 2002).

Conclusion

We show in this paper that the DE model both replicates human data on the categorical and coordinate dichotomy and doesn’t behave strictly in accordance with spatial frequencies. This calls into question the hypothesis about spatial frequencies driving coordinate and categorical lateralization. It also provides a point of differentiation between DFF theory and the DE model. Limited to spatial frequency information, DFF would be unable to replicate the conflicting data in Slotnick et al. (2001) as the DE model does.

In addition, we have shown in other experiments that DE models of a larger sigma encode more information at higher spatial frequencies, and vice-versa for smaller sigmas. Furthermore, stimulus size mediates this relationship: as image

size increases, so too does the point where the LH networks outperform the RH networks. We are pursuing this as an explanation for the relative frequency effect, noted in Christman, Kitterle, and Hellige (2001).

The DE model already accounts for faces and the categorical and coordinate results. If we can account for the relative frequency effect, we have superseded the DFF with a model that is biologically grounded and is informative about experiments to run in biology and psychology.

Acknowledgements

GWC is supported in part by NSF grant SMA 1041755 to the Temporal Dynamics of Learning Center and a gift from HP.

References

- Baker, D. P., Chabris, C. F., & Kosslyn, S. M. (1999). Encoding categorical and coordinate spatial relations without input-output correlations: New simulation models. *Cognitive science*, 23, 33–51.
- Christman, S., Kitterle, F. L., & Hellige, J. (1991). Hemispheric asymmetry in the processing of absolute versus relative spatial frequency. *Brain and Cognition*, 16, 62–73.
- Cipollini, B. (2014). *Modeling visual lateralization and interhemispheric communication*. Unpublished doctoral dissertation, University of California, San Diego.
- Gallagher, P., & Dagenbach, D. (2007). Manipulating noise frequencies alters hemispheric contributions to decision making. *Brain and cognition*, 64, 42–49.
- Galuske, R. A., Schlote, W., Bratzke, H., & Singer, W. (2000). Interhemispheric asymmetries of the modular structure in human temporal cortex. *Science*, 289, 1946–1949.
- Hellige, J. B., & Michimata, C. (1989). Categorization versus distance: Hemispheric differences for processing spatial information. *Memory & Cognition*, 17(6), 770–776.
- Hsiao, J. H., Cipollini, B., & Cottrell, G. W. (2013). Hemispheric asymmetry in perception: A differential encoding account. *Journal of Cognitive Neuroscience*, 25, 998–1007.
- Hsiao, J. H.-w., Shahbazi, R., & Cottrell, G. W. (2008). Hemispheric asymmetry in visual perception arises from differential encoding beyond the sensory level. In *Proceedings of the Cognitive Science Society* (Vol. 30).
- Ivry, R. B., & Robertson, L. C. (1998). *The two sides of perception*. Cambridge, MA: The MIT Press.
- Jager, G., & Postma, A. (2003). On the hemispheric specialization for categorical and coordinate spatial relations: A review of the current evidence. *Neuropsychologia*, 41, 504–515.
- 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. *Attention, Perception, & Psychophysics*, 48, 297–306.
- Kitterle, F. L., Hellige, J. B., & Christman, S. (1992). Visual hemispheric asymmetries depend on which spatial frequencies are task relevant. *Brain and cognition*, 20, 308–314.
- Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flel, A., . . . Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, 123, 2512–2518.
- Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological review*, 94, 148.
- Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., & Koenig, O. (1992). Categorical versus coordinate spatial relations: computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 562.
- Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. (1989). Evidence for two types of spatial representations: hemispheric specialization for categorical and coordinate relations. *Journal of experimental psychology: human perception and performance*, 15, 723.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive psychology*, 9, 353–383.
- Okubo, M., & Michimata, C. (2002). Hemispheric processing of categorical and coordinate spatial relations in the absence of low spatial frequencies. *Journal of cognitive neuroscience*, 14, 291–297.
- Rogers, L. J., & Andrew, R. (2002). *Comparative vertebrate lateralization*. Cambridge University Press.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.
- Sergent, J. (1982). The cerebral balance of power: confrontation or cooperation? *Journal of Experimental Psychology: Human Perception and Performance*, 8, 253.
- Sergent, J. (1985). Influence of task and input factors on hemispheric involvement in face processing. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 846.
- Sergent, J. (1991). Judgments of relative position and distance on representations of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 762.
- Slotnick, S. D., Moo, L. R., Tesoro, M. A., & Hart, J. (2001). Hemispheric asymmetry in categorical versus coordinate visuospatial processing revealed by temporary cortical deactivation. *Journal of Cognitive Neuroscience*, 13(8), 1088–1096.
- Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1), 1929–1958.
- Stephan, K. E., Marshall, J. C., Friston, K. J., Rowe, J. B., Ritzl, A., Zilles, K., & Fink, G. R. (2003). Lateralized cognitive processes and lateralized task control in the human brain. *Science*, 301(5631), 384–386.