

My previous research has both whetted my appetite for more and humbled me. I have first-hand experience with the challenges of research, as well as its rewards. In the process, I have developed an appreciation of the difficulty of formulating research questions that are both interesting and tractable and of seeing large, complex projects through to completion.

Most recently, during my Baggett Fellowship at the University of Maryland, I developed and investigated a sparse-coding model of how the primary auditory cortex might be adapted to the statistics of natural sounds, including speech. The first phase of my Plan of Research is based on this work. Sparse coding is a statistical technique that, like principal component analysis, finds an optimal linear transformation between an uninformative way of representing stimuli to a more informative way. In the example of visual images, each image can be represented by a vector, where each dimension represents the brightness of one image pixel. However, considering a large set of images, the brightness of each pixel is highly dependent on the brightness of its neighbors, and so there is redundancy in this representation. Linear statistical models remove this redundancy by finding a set of components (which can also be represented as images), such that each original image is represented by the weighted sum of components. These components thus capture the statistical structure of the images. Sparse coding finds the set of components that minimizes the number of components required to represent each individual image. As it turns out, these sparse components of natural images are indistinguishable from the receptive fields of single neurons in primary visual cortex (V1) [2]. Each V1 neuron's receptive field describes that unit's preferred stimulus, whose presence most reliably produces a spike.

While sparse coding was created as a model of visual perception, the same statistical techniques are equally applicable to any multidimensional stimulus, including sound. My advisor, Prof. William Idsardi, and I were interested in finding informative features of speech sounds in an unsupervised way, which sparse coding has the potential to do. Also, after a review of the relevant literature, I discovered that very little attention had been given to such statistical models of auditory perception. I applied a sparse coding model to speech sounds that had been filtered by a cochlear model (simulating the input that the primary auditory cortex—A1—receives). The components learned by the model were compared with actual A1 receptive fields, and found to be very similar, both in terms of frequency, bandwidth, and rate selectivity but also in terms of temporal symmetry [1] (an extremely restrictive property of A1 receptive fields [3]). These results (especially the similar degrees of temporal symmetry) show that the auditory system, like the visual system, is adapted to the statistics of the natural world through evolution, development, or some other mechanism.

While at Maryland I was also involved, to varying degrees, in a variety of other projects. I ran subjects in an MEG study of the difference between the perception of certain musical intervals embedded in speech and non-speech sounds, and I designed software to automate the pre-processing of video stimuli for a study on the effect on the intelligibility of sign-language of reversing small, local segments of videos, while preserving the global ordering of these segments. I also worked with Prof. Jordan Boyd-Graber on the development of a hierarchical Bayesian model which segments a string of phonetic segments in order to learn words and phonetic categories.

The first major component of my undergraduate research experience was a Cognitive Science honors thesis, which was advised by Prof. Safa Zaki, and formulated with the help of Prof. Robert Port at Indiana University over the summer of 2008, (supported by the Tyng Scholarship from Williams). This project investigated the role that speech-sound representations of various sizes play in the processing of spoken words and non-words. Having many common subsequences slows the

processing of words (since those words tend to sound like many other words, which compete for activation) but speeds the processing of nonwords. By comparing the range of qualitative patterns of behavior different models were capable of producing, my thesis concluded that large, sublexical representations were necessary to explain the behavioral data (at least within the modeling framework I used). For this project, I was awarded Highest Honors.

This result validates, using explicit computational modeling, the explanation that had been given at a theoretical level in the discussion of the behavioral data. I learned how challenging it is to make such a connection between behavioral data and computational modeling, even when the general framework has already been proposed at a high level. Implementing and evaluating the models was extremely challenging, and I furthermore had to integrate this technically demanding work into a larger theoretical context. Prof. Zaki was instrumental in accomplishing this; because she was not an expert on either the model I used or the sort of data I was applying it to I had to frame the project in broad terms. I thus learned the importance of getting input from non-specialists in focusing computational projects and making them broadly relevant.

As an undergraduate, I also worked with Prof. Zaki on her projects on the cognitive psychology of categorization, starting in the summer of 2006 when I was hired as a research assistant until I graduated in 2009. We worked closely together to conceptualize and design studies, which I had primary responsibility for implementing and running. I also was responsible for the implementation and application of non-standard computational models for analyzing the data we collected. We are currently writing up two of these projects for publication.

The first was an investigation of the source of apparent dissociation between simple, easily-verbalizable categorization tasks and more complicated tasks that require the integration of information from multiple stimulus dimensions for optimal performance. Dissociations between these tasks have been used to argue for the existence of multiple, functionally and neurally distinct categorization systems. However, the particular study we were investigating confounded type of task with different types of difficulty. Prof. Zaki and I designed an experiment which removed the dissociation by removing the confound.

The second project we worked on, which is still ongoing, was on the role of transformational knowledge on categorization. Previous work has shown that subjects are sensitive to coherent transformations associated with categories, and that presenting the same category exemplars in random order vs. as a coherent transformation changes subject's categorization judgements about novel test stimuli. As currently formulated, most modern models of categorization do not predict and cannot account for these results. We conducted a study which replicated and extended this work, and are currently investigating modeling frameworks which may be able to explain our results.

I have thus had a range of research experience, both individually and in collaboration, focusing on categorization, statistical models, and speech perception. This experience has prepared me to execute the project proposed in this application, and to have a productive research career in the cognitive sciences.

- [1] Kleinschmidt, D., 2010. Efficient coding of speech in the auditory cortex. *poster presented at the 2nd Annual Neurobiology of Language Conference.*
- [2] Olshausen, B.A. et al., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583) 607–9.
- [3] Simon, J.Z., et al., 2007. Temporal symmetry in primary auditory cortex: implications for cortical connectivity. *Neural computation*, 19(3) 583–638.