# A Unified Model of Categorical Effects in Consonant and Vowel Perception

**Yakov Kronrod (yakov@umd.edu)**
**Emily Coppess (ercoppess@umd.edu)**
**Naomi H. Feldman (nhf@umd.edu)**
Department of Linguistics, 1401 Marie Mount Hall
University of Maryland, College Park, MD 20742

## Abstract

Consonants and vowels differ in the extent to which they are perceived categorically. We use a Bayesian model of speech perception to explore factors that might cause this difference. Simulations show that perception of vowels, fricatives, and stop consonants can all be captured under a single model in which listeners use their knowledge of phonetic categories to infer the sound that a speaker intended. This suggests that the differences in the way we perceive vowels and consonants, when viewed at the computational level, can be explained as parametric variation within a single framework.

**Keywords:** perceptual magnet effect; categorical perception; Bayesian modeling; computational linguistics

Phonetic categories influence perception of speech sounds, with stimuli belonging to different categories being easier to discriminate than stimuli from a single category (Liberman, Harris, Hoffman, & Griffith, 1957; Kuhl, 1991). However, different types of sounds differ in the degree to which they are perceived categorically. At one end of the spectrum, perception of stop consonants is strongly categorical. Discrimination is little better than would be expected if listeners used only category assignments to distinguish sounds, and between-category differences are extremely pronounced (Liberman et al., 1957). At the other end of the spectrum, vowel perception is much more continuous, with some even arguing that vowels display no categorical effects at all (Fry, Abramson, Eimas, & Liberman, 1962).

Researchers have used various mechanisms underlying speech perception to explain these differences. For example, differences have been proposed to stem from the way each type of sound is stored in memory (Pisoni, 1973) and to be related to innate auditory discontinuities that seem to influence stop consonant perception (Pisoni, 1977; Eimas, Siqueland, Jusczyk, & Vigorito, 1971). However, the qualitative similarity of categorical effects in consonants and vowels suggests that in some ways these are also instances of the same phenomenon. This raises the possibility that perceptual differences among different classes of sounds are quantitative rather than qualitative.

In this paper we explore these similarities and differences at Marr's (1982) computational level, looking at the optimal solution to the problem of inferring speakers' intended productions from the available acoustic information. We adapt a Bayesian model from Feldman, Griffiths, and Morgan (2009), in which listeners use their knowledge of phonetic categories to recover the sound a speaker intended. We show that an extended version of this model can account for perceptual data from stop consonants and fricatives as well as vowels. This suggests that differences in the degree to which vowels and consonants are perceived categorically can be explained as parametric variations in a single underlying model.

Our paper is organized as follows. First, we review evidence concerning categorical effects in consonants and vowels, giving an overview of the types of explanations that have been proposed to account for these data. We then describe the model from Feldman et al. (2009) in detail, focusing on their results for vowel perception. In the subsequent section, we present simulations testing our extended model on two types of consonants, stop consonants and fricatives, to determine whether a model built for vowels can also account for patterns in consonant perception. We conclude by summarizing our findings and discussing implications for theories of speech perception.

## Categorical Effects in Speech Perception

Categorical perception in stop consonants was first described by Liberman et al. (1957) as consisting of a sharp change in the identification function between different consonants, as well as a peak in the discrimination function at the location of the identification boundary. The authors proposed a model in which participants used only category assignments to determine whether sounds were the same or different. If the sounds belonged to different categories, then participants would respond *different*; otherwise, they would respond *same*. By examining participants' identification functions, Liberman et al. could use this model to predict the extent to which participants should be able to discriminate each pair of sounds. Participants' actual discrimination performance exceeded the model's predictions only by a small amount, and the authors took this as evidence of a strong categorical component in stop consonant perception. Liberman et al.'s experiment focused on stop consonants that differed by place of articulation, but similar findings have been obtained along the voice onset time (VOT) dimension as well (Wood, 1976).

Descriptions of categorical effects in vowels have focused primarily on the *perceptual magnet effect* (Kuhl, 1991). This effect was originally proposed as a within-category phenomenon, characterized by sounds near category centers being more difficult to discriminate than sounds near category edges, with an accompanying correlation between goodness ratings and discriminability. There is disagreement over whether categorical perception and the perceptual magnet effect are separate phenomena or different variants of the same phenomenon (e.g. Lotto, Kluender, & Holt, 1998). Some characteristics of the perceptual magnet effect are similar to pure categorical effects, such as reduced discriminability near category centers. Data from Iverson and Kuhl (2000) suggested that discrimination peaks near category boundaries are separable from correlations of discrimination and goodness ratings, whereas more recent studies have found that these two effects cooccur (Tomaschek, Truckenbrodt, & Hertrich, 2011). Regardless of terminology, however, categorical effects in vowel perception are much weaker than those found in consonant perception.

In addition to stop consonants and vowels, it is natural to consider categorical perception of another major class of speech sounds, fricatives. In this paper, we consider categorical perception of sibilant fricatives. There has been some disagreement over the degree of categorical perception in fricatives in previous research. Repp (1981) showed that fricatives follow patterns similar to the categorical perception found for stop consonants. However, in the same study, a subset of participants seemed to have perception that was more continuous, which Repp attributed to a choice between two processing strategies, acoustic and phonetic. Others have found that fricative perception is much less categorical than stop consonants (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Healy & Repp, 1982; Repp, 1984). Another more recent study showed identification patterns consistent with categorical perception together with a neural signature indicative of something like perceptual warping near category centers (Lago, Kronrod, Scharinger, & Idsardi, 2010).

These data set up a continuum ranging from nearly completely categorical perception of stop consonants to much more continuous perception of vowels, with fricatives falling somewhere in between. However, this continuum is not as clear cut as it may seem, as neural and behavioral evidence suggests that listeners attend to phonetic detail when perceiving stop consonants (Pisoni & Lazarus, 1974; Blumstein, Myers, & Rissman, 2005), and the degree of categorical perception in both consonants and vowels can be influenced by task-related factors (Pisoni, 1975; Repp, Healy, & Crowder, 1979). Nevertheless, the differences between consonant and vowel perception are robust. In what follows, we use a model to account for the variability in these effects within a single framework, identifying aspects of category structure that might contribute to differences in categorical effects across consonants and vowels.

## Model Overview

Our model is an extension of the model from Feldman et al. (2009). The model assumes that listeners are trying to recover phonetic detail about the speaker's intended production as well as category information. It differs from traditional models of categorical perception in that it recognizes two different sources of within-category variability: meaningful variability (also referred to as category variance) and noise variance. The category variance $\sigma_c^2$ is assumed to arise from processes that yield information useful to listeners, such as coarticulatory effects that allow listeners to predict the identity of upcoming sounds (Gow, 2001). Once a speaker selects a target production, $T$, from the category, there is assumed to be additional articulatory, acoustic, and perceptual noise $\sigma_S^2$ that further distorts this sound. This process results in a speech sound $S$ that is heard by listeners.

Listeners are trying to infer the target production through the noisy speech signal. To do this, listeners can use their knowledge that speakers tend to produce sounds near category centers. Hence, they rely both on category knowledge and on acoustic cues to recover the phonetic detail of a speaker's target production. If listeners encounter little noise and the category allows a large amount of meaningful variability (e.g., coarticulation), then listeners attend more to acoustic detail in perceiving sounds; however, in situations with high noise and low meaningful variability, they rely more on their knowledge of phonetic categories. This relationship between category variance and noise plays an important role in determining the degree to which perception is biased toward category centers, and it thus has the potential to account for differences in the degree to which vowels and consonants are perceived categorically.

Feldman et al.'s (2009) original model relied on a simplifying assumption that all categories considered by a listener have equal category variance. While this assumption might be adequate for vowels, other sound categories do not necessarily reflect this simplification, particularly voiced and voiceless stop consonants which have been shown to have substantial differences in VOT variance (Lisker & Abramson, 1964). Hence, we extend the original model proposed by Feldman et al. (2009) to allow for unequal category variances. This section gives an overview of our extended model; full derivations are omitted due to space limitations, but are parallel to those in Feldman et al.

The model assumes phonetic categories are Gaussian distributions of sounds along the relevant auditory dimensions, so that a speaker's target production is normally distributed around the category mean, $T|c \sim N(\mu_c, \sigma_c^2)$. Noise in the speech signal causes the stimulus heard by listeners to be normally distributed around the target production, $S|T \sim N(T, \sigma_S^2)$. We can integrate over all possible target productions $T$ to get an expression relating the perceived sound directly to the underlying categories,

$$S|c \sim N(\mu_c, \sigma_c^2 + \sigma_S^2) \tag{1}$$

In identification tasks, listeners recover a category given the sound $S$, which corresponds to computing the posterior distribution over category membership $p(c|S)$ in the model. They can compute this by applying Bayes' rule,

$$p(c|S) = \frac{p(S|c)p(c)}{\sum_c p(S|c)p(c)} \tag{2}$$

If we limit ourselves to two categories but relax the assumption that these have equal category variances, we need two means ($\mu_{c_1}$ and $\mu_{c_2}$) and category variance parameters ($\sigma_{c_1}$ and $\sigma_{c_2}$). We derive the identification function by substituting

| **Simulation** | **Means** | | **Variances** | | | **Category:Noise Variance Ratio** |
|---|---|---|---|---|---|---|
| | $\mu_{c_1}$ | $\mu_{c_2}$ | $\sigma^2_{c_1}$ | $\sigma^2_{c_2}$ | $\sigma^2_S$ | |
| Vowels<br>(Equal Variance) | F1=224 Hz<br>F2=2413 Hz | F1=423 Hz<br>F2=1936 Hz | 5,873 | 5,873<br>(Mels) | 878 | 6.69 |
| Stop Consonants<br>(Unequal Variance) | 60 ms | -0.3 ms | 253.9 | 14<br>(ms) | 82.3 | /p/: 3.09, /b/: 0.17 |
| Fricatives<br>(Unequal Variance) | 19.0 Barks | 15.99 Barks | 0.5992 | 0.5772<br>(Barks) | 0.3098 | /s/: 1.93, /ʃ/: 1.86 |

Table 1: Best fitting model parameters for vowels (Feldman et al., 2009), stop consonants, and fricatives.

Equation 1 into Equation 2 and following a parallel derivation to that in Appendix B from Feldman et al. (2009), yielding

$$p(c_1|S) = \frac{1}{1+\sqrt{\frac{\sigma_1^2}{\sigma_2^2}} \times exp\frac{(\sigma_2^2-\sigma_1^2)S^2+2(\mu_{c_2}\sigma_1^2-\mu_{c_1}\sigma_2^2)S+(\mu_{c_1}^2\sigma_2^2-\mu_{c_2}^2\sigma_1^2)}{2\sigma_1^2\sigma_2^2}} \tag{3}$$

where $\sigma_1^2 = \sigma_{c_1}^2 + \sigma_S^2$ and $\sigma_2^2 = \sigma_{c_2}^2 + \sigma_S^2$.

The model assumes that listeners recover the phonetic detail of a speaker's target production in addition to category information when perceiving a speech sound, and that they use this information when performing a discrimination task. Perceiving phonetic detail corresponds to computing the posterior distribution on target productions, $p(T|S)$. Applying Bayes' rule, where the prior $p(T)$ is a mixture of Gaussians and the likelihood $p(S|T)$ is Gaussian, we obtain a posterior distribution whose form is a mixture of Gaussians and whose mean is

$$E[T|S] = \sum_c p(c|S)\frac{\sigma_c^2 S + \sigma_S^2 \mu_c}{\sigma_c^2 + \sigma_S^2} \tag{4}$$

(see Feldman et al., 2009 for a full derivation). Each category makes a contribution to this posterior mean with magnitude proportional to the posterior probability of the sound belonging to that category, $p(c|S)$. The specific contribution of each category is to bias perception toward the category mean. The strength of the bias is controlled by the relationship between parameters $\sigma_c^2$ and $\sigma_S^2$, which represent the amount of meaningful variability and the amount of noise. Notice that the contribution of the category mean, $\mu_c$, is weighted by the noise variance, $\sigma_S^2$. This means that when there is more noise, listeners will rely more on their underlying knowledge of the categories. In contrast, the acoustic information, $S$, is weighed by the meaningful variance parameter, $\sigma_c^2$, such that when there is a lot of meaningful variability in the underlying category, listeners will pay more attention to the acoustic data. It is this relationship that will be critical to modeling differences in perception between different categories of sounds.

Feldman et al. (2009) applied their model to vowel perception (continuum from /e/ to /i/), obtaining a close fit to the multidimensional scaling data from Iverson and Kuhl (1995) (Figure 1). However, in analyzing their own data from an AX discrimination experiment, the patterns they found suggested that multidimensional scaling was distorting the perceptual patterns, and that the noise parameter needed to capture experimental data directly was much lower than they initially found.[1] Thus, the "Vowels (Equal Variance)" section of Table 1 shows the values they derived on the basis of their experimental data. As might be expected for relatively continuous vowel perception, these parameters showed high meaningful category variance relative to noise variance, indicating that the bias toward category centers was small. We use these parameters as a baseline for comparison in our consonant simulations.

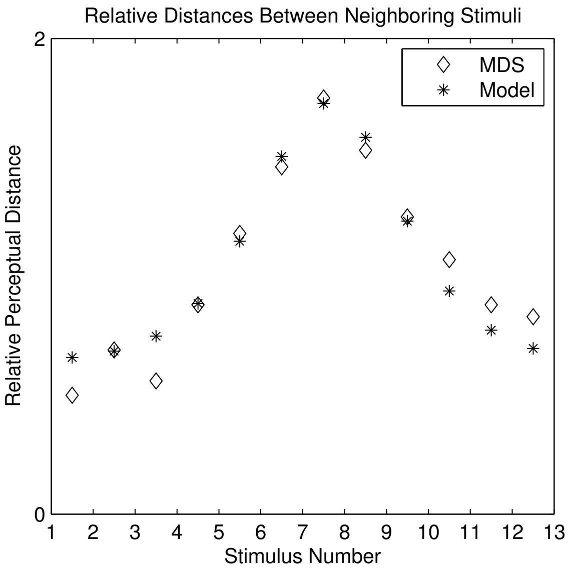

Figure 1: Figure from Feldman et al. (2009) showing interstimulus distances from Iverson and Kuhl's (1995) multidimensional scaling solution and the fitted model.

---

[1] In our simulations below, we select data that use the distance measure d' rather than multidimensional scaling data in order to avoid this type of discrepancy.

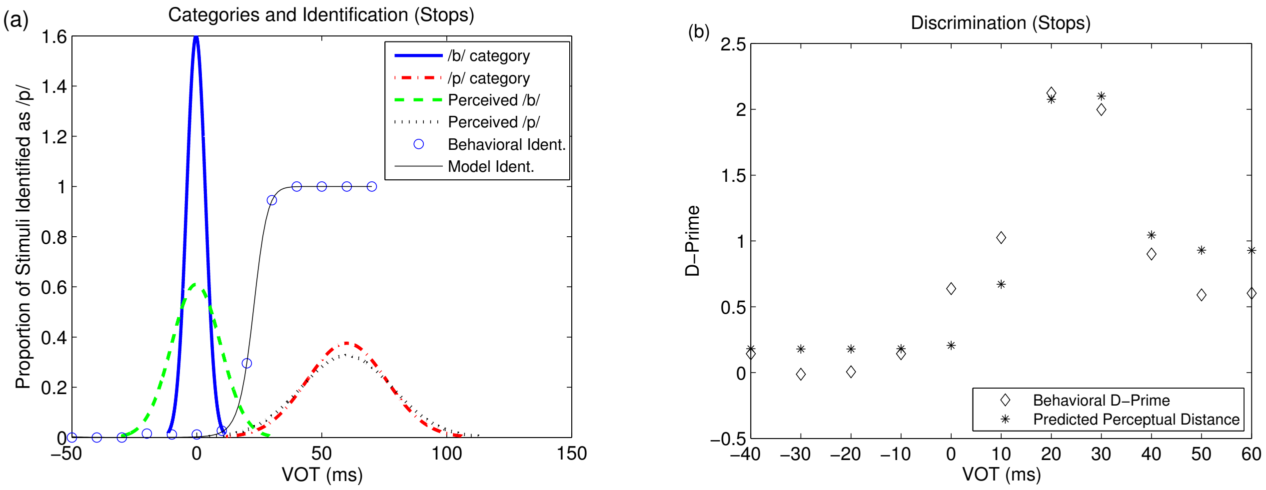

Figure 2: Stop consonants: (a) Underlying categories, perceived distributions, and identification curve in the model, together with behavioral identification data; (b) Interstimulus distances predicted by the model, together with behavioral d' data.

## Simulations

We applied this model to data from stop consonants and sibilant fricatives, deriving parameters on the basis of experimental data in order to determine whether categorical effects in each type of sound can be explained as the result of optimally inferring the phonetic detail of a speaker's target production. Examining the resulting parameters then allows us to assess the degree to which those parameters are adequate with regard to existing data, as well as examine the relationship between the two variance parameters and the degree of observed perceptual warping. Following Feldman et al. (2009), our general strategy for fitting our parameters was as follows:

1. Set $\mu_{c_1}$ on the basis of production data.
2. Determine $\mu_{c_2}$, $\sigma_1^2$, and $\sigma_2^2$ from identification data using Equation 3.
3. Determine the ratio of category variances, $\sigma_{c_1}^2$ and $\sigma_{c_2}^2$, to noise, $\sigma_S^2$, by fitting acoustic differences between percepts, $E[T|S]$, in the model (Equation 4) to a distance measure such as d'.

We need to set one of the means in order to obtain a single identifiable set of parameters. The model is then fit to identification data, allowing us to derive the other mean as well as both sums of variances (one corresponding to each category). Note that the only parameter being fit directly to the discrimination data is the ratio of meaningful category variance to noise variance, which is the parameter of interest for examining the degree of bias toward category centers exhibited by each class of sounds. In effect, the discrimination data provide a general test of the model's fit to behavioral data from each class of sounds.

## Stop Consonants

We first consider behavioral data from identification and discrimination experiments on stop consonants, which have been found to exhibit much stronger categorical effects in perception than vowels. Under our account, this difference might stem from low category variance relative to noise variance, such that listeners rely more on category information. If we are able to account for stop consonant perception with our model, then that would suggest that it may not be necessary to posit qualitative differences in the types of computations performed by listeners when perceiving consonants and vowels. It is not obvious that our model should be able to explain stop consonant data, however, as other factors such as innate phonetic boundaries (Eimas et al., 1971) or auditory discontinuities (Pisoni, 1977) might retain their influence on stop consonant perception even after phonetic learning is complete.

For this simulation we used identification and discrimination data from Wood (1976), who examined perception of /p/ and /b/ along a voice onset time (VOT) continuum. The continuum consisted of synthetic stimuli ranging from -50 to +70 ms VOT. A forced identification task as well as both a 10-ms and 20-ms difference AX discrimination task were administered. We used 20-ms discrimination data for our simulations. On the basis of data from Lisker and Abramson (1964), we set $\mu_{/p/}$ at 60 ms VOT and derived the remaining parameters from the identification and discrimination data. The identification fit produced an estimated value of -0.3 ms for the mean $\mu_{/b/}$, which was a close match to production data found in Lisker and Abramson (1964). The full set of parameters is found in section "Stop Consonants (Unequal Variance)" of Table 1, and the resulting identification curve and category distributions are shown in Figure 2(a). The fit between model and data is very close: the model is even able to predict the reduced within-category discriminability of voiced stops compared to voiceless stops that is observed in the empirical data.

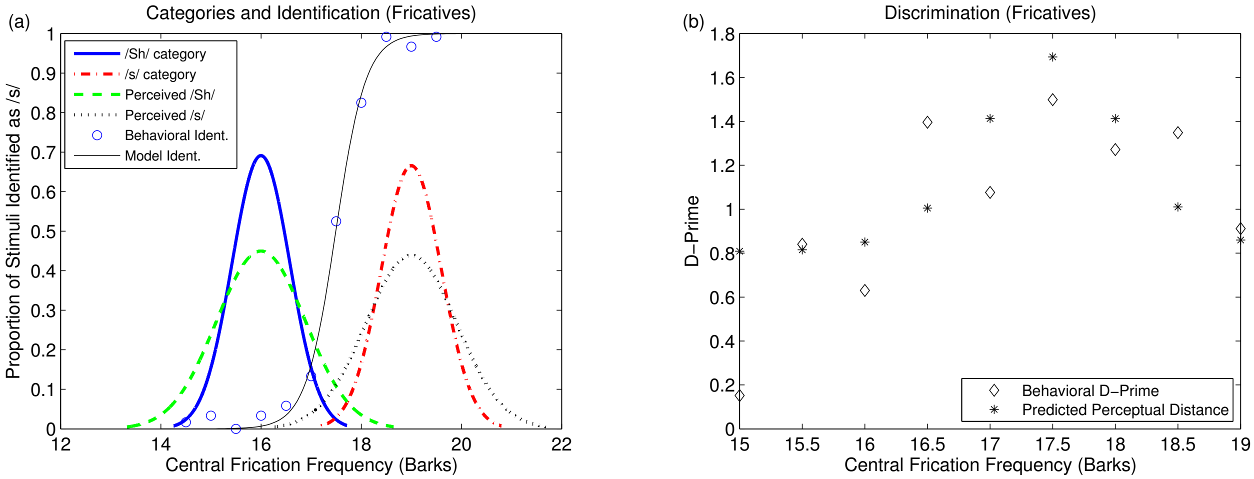

Figure 3: Sibilant fricatives: (a) Underlying categories, perceived distributions, and identification curve in the model, together with behavioral identification data; (b) Interstimulus distances predicted by the model, together with behavioral d' data.

This can be seen in Figure 2(b), where the perceptual distance between stimuli toward the left side of the continuum is lower than that toward the right side of the continuum.

As predicted, the ratio of category variance to speech signal noise was lower than that obtained for vowels for both categories of stop consonants, voiced and voiceless. These findings suggest that stop consonants have less meaningful within-category variance relative to noise variance than vowels, leading the listener to rely on prior category knowledge in inferring the speakers' target production. This causes a greater pull toward category centers and hence stronger categorical perception. Perceptual bias is particularly strong in voiced stops due to their low category variance.

## Sibilant Fricatives

The previous simulation indicates that the model provides a good account for stop consonants as well as vowels. We next apply our model to sibilant fricatives. Sibilant fricatives are obstruents (like stop consonants), but their characteristic noise components at higher frication frequencies show some similarity to the higher formant structures of vowels. As discussed above, there has been conflicting evidence on the strength of categorical effects they exhibit. These factors make fricatives an interesting modeling target.

For this simulation we used identification and discrimination data along the /s/-/ʃ/ continuum from Lago et al. (2010). The continuum consisted of 11 tokens with central frication frequencies varying from 14.5 to 19.5 Barks. A forced identification task as well as a 2-step AX discrimination task were administered to 12 participants. For our model, we fixed the value of $\mu_{/s/}$ to 19.0 Barks based on natural productions by an adult male participant and derived values for the remaining parameters by fitting the model to behavioral identification and discrimination data. The resulting parameter values are given in the "Fricatives (Unequal Variance)" entry in Table 1.

Figure 3(a) shows the identification data and the identification curve used in our model, together with the underlying and perceived category distributions that correspond to the parameters used in our simulation. Figure 3(b) compares the model predictions to the observed discrimination measures. The fit is not perfect, due in part to noisy data from the original experiment, but both data and model show the peak in discrimination at the same location as the inflection point in the identification data.

Given that fricatives tend to be perceived more categorically than vowels but less so than consonants, we might expect the category variance to noise ratio to be smaller for fricatives than for vowels, leading to a larger perceptual bias toward category centers, but larger than that for stop consonants, indicating more attention focused on acoustic cues. As predicted, the ratios for the sibilant fricatives are reduced compared to the parameters estimated for vowels (1.93 and 1.86 compared to 6.69). Additionally, we see that they are close to the ratio for the voiceless stops but much higher than that of the voiced consonants, suggesting that they may be closer to the stop consonant end of the spectrum in terms of their degree of bias toward category centers.

## Discussion

This paper used a Bayesian model to investigate the relationship between categorical effects in consonant and vowel perception. Our results suggest that these effects can be explained at Marr's (1982) computational level by the same underlying principles: Listeners use their knowledge of phonetic categories to optimally infer a speaker's target production through a noisy speech signal, and this causes their perception to be biased toward category centers. The model accounts for differences in the strength of categorical effects by assigning consonants less meaningful variability, compared with noise variance, than vowels.

Our analysis is reminiscent of an analysis pursued by Pisoni (1973), Liberman et al. (1967), and others. Their account proposes that speech perception incorporates a phonetic mode of perception, i.e., categorical perception, and an auditory mode of perception, i.e., continuous perception. Pisoni (1973) argued that differences between consonant discriminability and vowel discriminability could be accounted for by assuming that listeners have less access to auditory short-term memory when hearing consonant stimuli than when hearing vowel stimuli. This distinction between phonetic and acoustic modes of perception is parallel to the weighted average in Equation 4, where acoustic information is weighted by the category variance and the category mean is weighted by the noise variance. When noise variance dominates over category variance, listeners rely more on the category mean rather than acoustic information (i.e. phonetic mode). Otherwise, acoustic information receives more weight and within-category discriminability increases (i.e. auditory mode). Looking at ratios of category variance to noise variance across consonants and vowels we see that for vowels category variance exerts much more influence than noise variance and therefore listeners' perception is drawn less towards the category center, causing within-category discriminability to increase (i.e. continuous perception). For consonants the ratio is smaller, coinciding with a decrease in within-category discriminability (i.e. categorical perception).

Our findings suggest that perception of three types of sounds – vowels, stop consonants, and fricatives – adheres to the same abstract computational principles. Importantly, however, the idea that listeners are performing the same computation at an abstract level does not necessarily mean that the underlying mechanisms are identical. Our analysis simply suggests that perception of each type of sound has been optimized to allow listeners to recover the sound intended by a speaker. Bias toward category centers may be implemented differently across different classes of sounds, and separate mechanisms are almost certainly necessary for extracting the various cues we have used as input to our model (formant frequencies for vowels; voice onset time for stop consonants; and central frication frequencies for fricatives). In future work we hope to explore these issues by considering perception of a fourth class of speech sounds, nasals, and by linking our computational approach with descriptions of sound perception at the algorithmic and implementational levels.

**Acknowledgments** We thank Sol Lago, Mathias Scharinger, and Bill Idsardi for sharing behavioral data from their experiments with fricatives. We also thank the Computational Psycholinguistics group, the PFNA Sounds group, and the Language Science Lunch group for valuable discussion and feedback. Finally, we thank Bill Idsardi and four anonymous reviewers for their insightful commentary and suggestions on earlier versions of this paper. This work was supported in part by NSF IGERT grant DGE-0801465.

## References

Blumstein, S. E., Myers, E. B., & Rissman, J. (2005). The perception of voice onset time: An fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience*, *17*(9), 1353-1366.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303-306.

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review*, *116*(4), 752-782.

Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, *5*, 171-189.

Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, *45*, 133-159.

Healy, A. F., & Repp, B. H. (1982). Context independence and phonetic meditation in categorical perception. *Journal of Experimental Psychology: Human Perception and Performance*, *8 (1)*, 68-80.

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America*, *97*(1), 553-562.

Iverson, P., & Kuhl, P. K. (2000). Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism? *Perception and Psychophysics*, *62*(4), 874-886.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, *50*(2), 93-107.

Lago, S., Kronrod, Y., Scharinger, M., & Idsardi, B. (2010). *Categorical perception of [s] and [sh]: An MMN study.* Neurobiology of Language Conference. San Diego, CA.

Liberman, A. M., Cooper, F., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431-461.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*(5), 358-368.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384-422.

Lotto, A. J., Kluender, K. R., & Holt, L. L. (1998). Depolarizing the perceptual magnet effect. *Journal of the Acoustical Society of America*, *103*(6), 3648-3655.

Marr, D. (1982). *Vision: A computational investigation in the human representation of visual information*. San Francisco: Freeman & Co.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, *13*(2), 253-260.

Pisoni, D. B. (1975). Auditory short-term memory and vowel perception. *Memory and Cognition*, *3*(1), 7-18.

Pisoni, D. B. (1977). Identification and discrimination of the relative onset time of two component tones: Implications for voicing perception. *Journal of the Acoustical Society of America*, *61*(5), 1352-1361.

Pisoni, D. B., & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *Journal of the Acoustical Society of America*, *55*(2), 328-333.

Repp, B. H. (1981). Two strategies in fricative discrimination. *Perception and Psychophysics*, *30 (3)*, 217-227.

Repp, B. H. (1984). Categorical perception: Issues, methods, findings. *Speech and Language: Advances in Basic Research and Practice*, *10*, 243-335.

Repp, B. H., Healy, A. F., & Crowder, R. G. (1979). Categories and context in the perception of isolated steady-state vowels. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(1), 129-145.

Tomaschek, F., Truckenbrodt, H., & Hertrich, I. (2011). Processing german vowel quantity: Categorical perception or perceptual magnet effect? *Proceedings of the 17th International Conference of Phonetic Sciences*, 2002-2005.

Wood, C. C. (1976). Discriminability, response bias, and phoneme categories in discrimination of voice onset time. *Journal of the Acoustic Society of America*, *60 (6)*, 1381-1389.