A Rational Account of the Perceptual Magnet Effect

Naomi H. Feldman (naomi_feldman@brown.edu)
Department of Cognitive and Linguistic Sciences, Brown University, Providence, RI 02912 USA

Thomas L. Griffiths (tom_griffiths@berkeley.edu)
Department of Psychology, University of California at Berkeley, Berkeley, CA 94720 USA

Abstract

The perceptual magnet effect involves reduced discriminability near prototypical vowel sounds in the native language. We present a Bayesian model to explain why this reduced discriminability might occur: it arises as a consequence of optimally solving the statistical problem of perceiving speech sounds in the presence of noise. In the optimal solution to this problem, listeners' perception of speech sounds is biased toward the means of phonetic categories because they use knowledge of these categories to guide their inferences about speakers' target productions. Simulations show that the predictions of the model closely correspond to human data.

Keywords: perceptual magnet effect; speech perception

It has long been known that categories influence perception, especially in the domain of speech sounds (Liberman, Harris, Hoffman, & Griffith, 1957). Similar categorical effects have been described in other domains, including color perception (Davidoff, Davies, & Roberson, 1999) and artificial categories of objects (Goldstone, Lippa, & Shiffrin, 2001). Despite widespread interest in this phenomenon, the reasons and mechanisms behind the connection between categories and perception remain unclear.

Categorical perception, the extreme case in which subjects can discriminate two speech sounds only when they belong to different phonetic categories, has been observed primarily for consonants. The role of phonetic categories in the perception of vowels has been more controversial. Fry, Abramson, Eimas, and Liberman (1962) first noted that vowel perception was much more continuous than consonant perception. More recently, however, Kuhl and colleagues have found evidence of poor discrimination near phonetic category prototypes, a phenomenon they have called the *perceptual magnet effect* based on the idea that native language prototypes pull neighboring speech sounds toward them (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; but see Lotto, Kluender, & Holt, 1998, for an alternative point of view). The effect has been demonstrated in the English /i/ category (Iverson & Kuhl, 1995), the German /i/ category (Diesch, Iverson, Kettermann, & Siebert, 1999), and the Swedish /y/ category (Kuhl et al., 1992). Though the effect remains elusive in other English vowels (Thyer, Hickson, & Dodd, 2000), language-specific shrinking of perceptual space has also been shown near the English /r/ and /l/ prototypes in English but not Japanese speakers (Iverson & Kuhl, 1996; Iverson et al., 2003).

While it has been argued that the perceptual magnet effect is a separate phenomenon from categorical perception (Iverson & Kuhl, 2000), most evidence for the phenomenon has shown characteristics that are qualitatively similar to other types of categorical effects: perceptual space is shrunk near the centers of categories and expanded near category boundaries. An explanation for the perceptual magnet effect, then, should account for both of these components while still allowing for continuous vowel perception in which subjects can discriminate within-category contrasts.

Previous computational models of the perceptual magnet effect have attributed various roles to phonetic categories. Guenther and Gjaja (1996) proposed a neural network model in which shrinkage of perceptual space near category centers emerges through an unsupervised learning mechanism trained on specific distributions of speech sounds, with categories playing no explicit role. Similarly, Damper and Harnad (2000) have argued based on neural models that categorical perception is an emergent property of the stimulus continuum. At the other extreme, Lacerda (1995) has proposed a model in which the perceptual magnet effect emerges as a side-effect of a classification problem. The goal of listeners is to classify sounds into phonetic categories; perception involves retrieving a set of numerical values that reflect the sound's similarity to each phonetic category in a language.

In this paper we take a novel approach to modeling the perceptual magnet effect. In the spirit of Marr (1982) and Anderson (1990), we consider the abstract computational problem posed by speech perception and show that the perceptual magnet effect emerges as part of the optimal solution to this problem. Specifically, we assume that listeners are optimally solving the problem of perceiving speech sounds in the presence of noise. Listeners have knowledge of discrete phonetic categories, but their goal in speech perception is to extract phonetic detail in addition to category membership in order to reconstruct coarticulatory and non-linguistic information. This is a difficult problem for listeners because they cannot hear the speaker's target production directly. Instead, they hear speech sounds that are similar to the speaker's target production but that have been altered through articulatory and acoustic noise. We formalize this problem using Bayesian statistics and show that the optimal statistical solution to this problem produces the perceptual magnet effect.

The paper is organized as follows. In the next section, we introduce our model of speech perception. The following section explores the relationship between category membership, perceptual bias, and perceptual distance, laying the background for simulations that provide a quantitative comparison between the model's behavior and empirical data. Finally, the discussion revisits the model's assumptions and draws parallels to previous models of the perceptual magnet effect.

Bayesian Model of Speech Perception

Our model sets up perception of speech sounds as a statistical problem. The goal of listeners, in perceiving a speech sound, is to reconstruct the acoustic detail of a speaker's target production. They extract this detail using the information that is available to them from the speech signal and their prior knowledge of phonetic categories. Phonetic categories are defined in this model as Gaussian distributions of speech sounds; in producing a speech sound, speakers select a phonetic category and articulate a target production from that category. Listeners hear a distorted version of this target production due to articulatory and acoustic noise, approximated in the model as Gaussian noise. In laying out the mathematics of the model, we begin by examining the case of a hypothetical language with one phonetic category; we then move on to the more complex case of multiple categories.

One Phonetic Category

When listeners perceive a speech sound, they can assume it was generated by selecting a target production from a phonetic category and then generating a noisy speech sound based on the target production. More formally, if phonetic category c has mean μ_c and variance σ_c^2, speakers generate target production T from that phonetic category. Listeners hear speech sound S through speech signal noise σ_S^2. This statistical model can be written as

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

$$S|T \sim N(T, \sigma_S^2) \tag{2}$$

Listeners hear the speech sound S and know the structure and location of phonetic categories in their native language; their task is to infer the speaker's target production T based on this information.

Using the speech sound S as data and the structure of phonetic category c as a prior, listeners can use Bayes' rule

$$p(T|S,c) \propto p(S|T)p(T|c) \tag{3}$$

to infer the speaker's target production T. The likelihood $p(S|T)$, given by the speech signal noise (Equation 2), assigns highest probability to speech sound S; the prior $p(T|c)$, given by phonetic category structure (Equation 1), assigns highest probability to the mean of the phonetic category. Since both likelihood and prior are Gaussian, their combination yields a posterior distribution that is a Gaussian whose mean falls between the speech sound S and the mean μ_c of the phonetic category. This posterior probability distribution can be summarized by its mean (the expectation of T given S and c), which is

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

The optimal guess at the speaker's intended production, then, is a weighted average of the speech sound heard and the mean of the speech sound's phonetic category, where the weights are determined by the ratio of category variance to speech signal noise.[1]

Equation 4 formalizes the idea of a perceptual magnet: the term μ_c pulls the perception of speech sounds toward the mean of the phonetic category, effectively shrinking perceptual space around the phonetic category. The resulting perceptual pattern is shown in Figure 1 (a). Note that if there is no uncertainty about category membership, perception of speech sounds further from the category mean is more biased than perception of speech sounds closer to the category mean. Consequently, all of perceptual space is shrunk to the same degree. If listeners are certain that all sounds belong to a single category, perceptual bias toward the category mean causes all of perceptual space to shrink toward the center of the category.

The analysis given above is the solution to a standard problem in Bayesian statistics (e.g., Gelman, Carlin, Stern, & Rubin, 1995), but Huttenlocher, Hedges, and Vevea (2000) also worked out the solution to an inference problem similar to this in the domain of non-linguistic stimuli. They noted that subjects' responses in visual stimulus reproduction tasks are generally biased toward the mean of the set of stimuli in an experiment and developed a model to account for that bias. Their model of visual stimulus reproduction assumes that subjects in an experiment form an implicit category consisting of all the stimuli they have seen and that they use this implicit category to correct for memory uncertainty when asked to reproduce a stimulus. For a Gaussian category distribution and Gaussian noise, the optimal way to correct for memory uncertainty using this implicit category is to bias all responses toward the mean value of the category, which in this case is the mean value of the set of stimuli. The mathematical analysis of this problem is nearly identical to ours, reflecting the similar structure of the two problems.

Multiple Phonetic Categories

The one-category case, while appropriate to explain the bias caused by an implicit category of visual stimuli within an experimental setting, is not appropriate for describing natural language. We therefore extend the model so that it applies to the more realistic case of multiple phonetic categories. With multiple categories, the probability that a particular category generated a speech sound can be calculated using Bayes' rule:

$$p(c|S) = \frac{p(S|c)p(c)}{\sum_c p(S|c)p(c)} \tag{5}$$

where $p(S|c)$ is computed by summing over all possible target sounds, $p(S|c) = \int p(S|T)p(T|c)\; dT$, and $p(c)$ reflects the prior probability assigned to category c.

The probability that a particular category generated a speech sound can be used in evaluating what the speaker's target production might have been. In reconstructing the target, listeners should take into account all the categories that could

[1] The expectation is optimal when the penalty for misidentifying a speech sound increases with squared distance from the target.

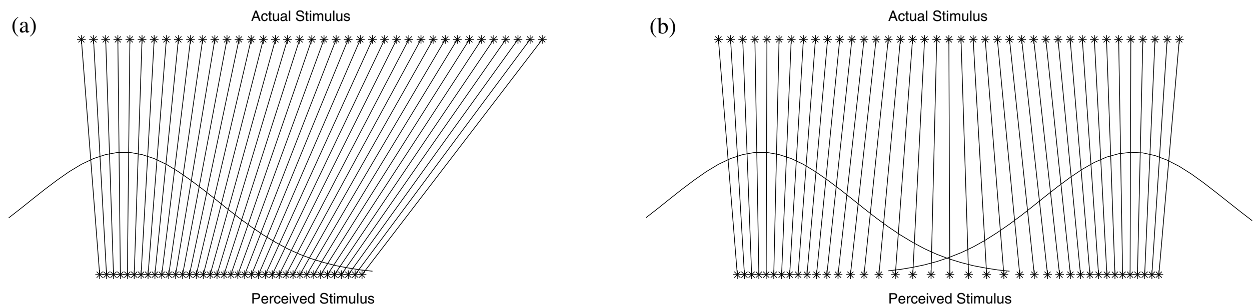

Figure 1: Predicted relationship between acoustic and perceptual space in the case of (a) one category and (b) two categories.

have produced the speech sound they heard, but they should weight the influence of each category by the probability that it produced the speech sound. To do this, they marginalize over phonetic categories, so that

$$p(T|S) = \Sigma_c p(T|S,c)p(c|S) \tag{6}$$

where $p(T|S,c)$ is the posterior distribution over T computed by assuming that it comes from category c, as in the single category case analyzed above (Equation 3).

The posterior distribution on T given S is now a mixture of Gaussians rather than a single Gaussian, but we can still compute its mean. Restricting our analysis to the case of categories with equal variance, the expected value of T given S, aggregating over all categories, is simply

$$E[T|S] = \frac{\sigma_c^2}{\sigma_c^2+\sigma_S^2}S + \frac{\sigma_S^2}{\sigma_c^2+\sigma_S^2}\Sigma_c p(c|S)\mu_c \tag{7}$$

The estimated value of T is thus a weighted average of speech sound S and the means μ_c of all the phonetic categories that might have produced S, where the contribution of μ_c is regulated by $p(c|S)$. When listeners are certain of a speech sound's phonetic category, this reduces to Equation 4, and perception of a speech sound S is biased toward the mean of its phonetic category. However, a speech sound directly on the border between two categories, with a high probability of having been generated from either, is pulled simultaneously toward both category means, each cancelling out the other's effect. Shrinkage of perceptual space is thus strongest in areas of unambiguous speech sound categorization – the centers of phonetic categories – and weakest at category borders. The correspondence between acoustic and perceptual spaces for the two-category case is shown in Figure 1 (b).

Characterizing Perceptual Warping

Our statistical analysis of the problem of speech perception establishes a simple function mapping an acoustic stimulus, S, to a percept of the intended speech sound, given by $E[T|S]$. In the case where multiple phonetic categories are present, this mapping is given by Equation 7. In order to formally analyze the qualitative behavior of the model, this section focuses on the relationship between three measures in the two-category case: *identification*, the posterior probability of category membership; *displacement*, the difference between the actual and perceived stimulus; and *warping*, the degree of shrinkage or expansion of perceptual space.

In the two-category case, under the assumptions outlined above, the identification function has the form of a logistic function. If both categories have equal prior probability, the posterior probability of membership in a given category c_1 can be written as

$$p(c_1|S) = \frac{1}{1+e^{-gS+b}} \tag{8}$$

where $g = \frac{\mu_1-\mu_2}{\sigma_c^2+\sigma_S^2}$ and $b = \frac{\mu_1^2-\mu_2^2}{2(\sigma_c^2+\sigma_S^2)}$. A logistic function of this form is shown in Figure 2 (a). In areas of certain categorization, the identification function is at either 1 or 0; a value of 0.5 indicates maximum uncertainty about category membership.

Displacement involves a comparison between the location of a speech sound in perceptual space $E[T|S]$ and its location in acoustic space S, where

$$E[T|S] - S = \frac{\sigma_S^2}{\sigma_c^2+\sigma_S^2}\left(\sum_c p(c|S)\mu_c - S\right) \tag{9}$$

In the one-category case, this means the amount of displacement is proportional to the distance between the speech sound S and the mean μ_c of the phonetic category. As speech sounds get farther away from the category mean, they are pulled proportionately farther toward the center of the category. The dashed lines in Figure 2 (b) show two cases of this. In the case of multiple categories, the amount of displacement is proportional to the distance between S and a weighted average of the means of more than one phonetic category. This is shown in the solid line, where ambiguous speech sounds are displaced less than would be predicted in the one-category case because of the competing influence of a second category mean.

Finally, perceptual warping can be characterized based on the distance between two neighboring points in perceptual space that are separated by a fixed step ΔS in acoustic space. This quantity is reflected in the distance between neighboring points on the bottom layer of each diagram in Figure 1. By the standard definition of the derivative as a limit, as ΔS approaches zero this measure of perceptual warping corresponds to the derivative of $E[T|S]$ with respect to S. This

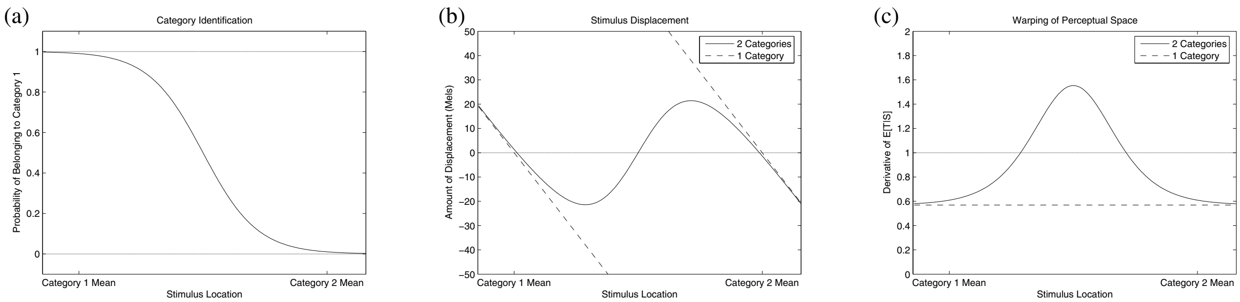

Figure 2: Measures of (a) identification, (b) displacement, and (c) warping. Solid lines show the values when both categories are considered; dotted lines show corresponding values in a hypothetical language with only a single category.

derivative is

$$\frac{dE[T|S]}{dS} = \frac{\sigma_c^2}{\sigma_c^2+\sigma_S^2} + \frac{\sigma_S^2}{\sigma_c^2+\sigma_S^2}\sum_c \mu_c \frac{dp(c|S)}{dS} \tag{10}$$

where the last term is straightforward to compute, being the derivative of the logistic function given in Equation 8.

When the derivative given in Equation 10 has a value greater than one, perceptual space is expanded relative to acoustic space; a value of less than one indicates shrinkage of perceptual space. This equation demonstrates that distance between two neighboring points in perceptual space is a linear function of the rate of change of $p(c|S)$, the identification function. The identification function is changing most rapidly near category boundaries, in areas of highest uncertainty, resulting in greater perceptual distances between neighboring stimuli near the edges of phonetic categories. In the one category case, shown by the dotted line in Figure 2 (c), the identification function is constant, so the warping function is always less than one and all of perceptual space is shrunk. The two category case, shown by the solid line, includes a portion of expanded perceptual space in the area where the identification function is changing most rapidly.

Taken together, these three measures show that interaction between neighboring phonetic categories produces a pattern of perceptual warping in which speech sounds near a category mean are extremely close together in perceptual space, whereas speech sounds near the edges of a category are much farther apart. This perceptual pattern results from a combination of two factors, both of which were suggested by Liberman et al. (1957) in reference to categorical perception. The first is acquired similarity within categories due to perceptual bias toward category means; the second is acquired distinctiveness between categories due to the presence of multiple categories. Under the assumptions of this model, then, the optimal solution for a rational perceiver is to shrink perceptual space near phonetic category centers and expand perceptual space near category boundaries. The pattern of warping found in the perceptual magnet effect falls neatly out of an analysis in which listeners use knowledge about the distribution of speech sounds in phonetic categories to optimally infer phonetic detail in the presence of speech signal noise.

Simulations

The formal results presented in the previous section establish that the qualitative predictions of our Bayesian model are broadly compatible with the warping associated with the perceptual magnet effect. In this section, we present a quantitative test of the model, examining whether a reasonable set of parameters can be found to match empirical data.

Some of the most detailed quantitative evidence for the perceptual magnet effect comes from a study by Iverson and Kuhl (1995), who used signal detection theory and multidimensional scaling to map perceptual distances near prototypical and non-prototypical /i/ vowels. They tested adults' discrimination of 13 stimuli along a single vector in F1-F2 space, ranging from F1 of 197 Hz and F2 of 2489 Hz (classified as /i/) to F1 of 429 Hz and F2 of 1925 Hz (classified as /e/) in 30-mel[2] intervals. Their multidimensional scaling results, shown in Figure 3, were used to test the model quantitatively.

Parameters in the model were based as much as possible on empirical measures in order to reduce the number of free parameters. The parameters that needed to be specified were $\mu_{/i/}$, the /i/ category mean; $\mu_{/e/}$, the /e/ category mean; σ_c^2, the category variance; and σ_S^2, the speech signal noise.

Subjects' goodness ratings from Iverson and Kuhl (1995) were first used to specify $\mu_{/i/}$. The mean of the /e/ category, $\mu_{/e/}$, and the sum of the variances, $\sigma_c^2+\sigma_S^2$, were calculated based on the gain and bias of a logistic function that was fit to the phoneme identification curves from Lotto et al. (1998).[3] The ratio between the category variance σ_c^2 and the speech signal noise σ_S^2 was the only remaining free parameter, and its value was chosen in order to maximize the fit to Iverson and Kuhl (1995)'s multidimensional scaling data.

A direct comparison was made by calculating the expectation $E[T|S]$ for each of the 13 stimuli according to Equation 7 and then determining the distance in mels between the expected values of neighboring stimuli. These distances were compared with the distances between stimuli in the multi-

[2] The mel frequency scale equates psychophysical distance.

[3] Because the stimuli in the MDS task were presented to subjects in all possible pairings, we averaged the identification curves obtained with prototype and nonprototype referents to produce a single intermediate identification curve.

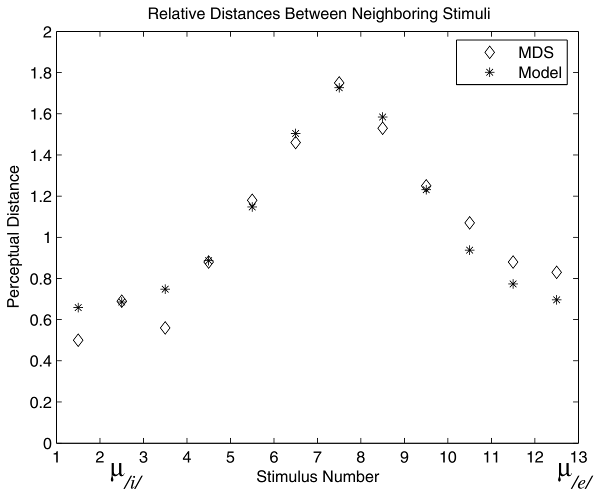

Figure 3: Relative distances between neighboring stimuli in Iverson and Kuhl (1995)'s multidimensional scaling analysis and in the model. The labels $\mu_{/i/}$ and $\mu_{/e/}$ show locations of category means in the model; distances are smallest near these means and largest at the boundary between them.

dimensional scaling solution. Since multidimensional scaling gives relative, and not absolute, distances between stimuli, this comparison was evaluated based on whether mel distances in the model were proportional to distances found through multidimensional scaling. As shown in Figure 3, the model yielded an extremely close fit to the empirical data, with interstimulus distances that were proportional to those found in multidimensional scaling (r=0.97). This simulation used the following parameters:

$\mu_{/i/}$: F1=224 Hz, F2=2,413 Hz
$\mu_{/e/}$: F1=423 Hz, F2=1,936 Hz
σ_c^2: 5,873 (σ_c = 77 mels)
σ_S^2: 4,443 (σ_S = 67 mels)

The fit obtained between the simulation and the empirical data is extremely close; however, model parameters derived in this simulation are meant to serve only as a first approximation of the actual parameters in vowel perception. Because of the variability that has been found in subjects' goodness ratings of speech stimuli, it is likely that these parameters are somewhat off from their actual values, and it is also possible that the parameters vary between subjects.

To understand the behavior of the model under various parameter combinations, we varied the prior probability, category variance, and speech signal noise independently in simulations. Varying the prior probability of the categories causes a shift in the discriminative boundary between the /i/ and /e/ categories. The boundary is shifted toward the category with lower prior probability, so that a larger region of acoustic space between the two categories is classified as belonging to the category with higher prior probability. This sort of boundary shift has been documented based on lexical context (Ganong, 1980): in contexts where one phoneme would form a lexical item and the other would not, phoneme boundaries are shifted toward the phoneme that makes the non-word.

Manipulating category variance yields extreme categorical perception in categories with low variance and perception that is less categorical in categories with high variance. When the variance is so high that the distribution of speech sounds in the two categories is unimodal, the model predicts that all speech sounds are biased toward a point between the two category means.

Finally, manipulating the speech signal noise produces a complex effect. Whereas adding low levels of noise makes perception more categorical, there comes a point where noise is too high to determine which category produced a speech sound, blurring the boundary between categories.

In this section, we have demonstrated through quantitative simulations based on a reasonable set of parameters that the model can reproduce Iverson and Kuhl (1995)'s quantitative multidimensional scaling data for the /i/ and /e/ categories. In addition, we have shown that the model captures similar patterns of perception using a wide range of parameter values and that parameter changes cause predictable shifts in boundary location and in the degree to which perception is categorical. Our model thus provides quantitative, as well as qualitative, predictions of the perceptual magnet effect.

Discussion

This paper has described a Bayesian model of speech perception in which listeners reconstruct the acoustic detail of a speaker's target production based on the speech sound they hear and their prior knowledge of phonetic categories. Uncertainty in the speech signal causes listeners to infer a target production that is closer to the mean of a phonetic category than the speech sound they actually heard. Assuming a language has multiple phonetic categories, listeners must first infer which category produced a speech sound and can then use that information to guide their inference of acoustic detail.

A basic assumption in the model is that listeners have knowledge of phonetic categories but are trying to infer phonetic detail. This assumption contrasts with previous models but is consistent with empirical data showing that listeners are sensitive to sub-phonemic detail at both neural and behavioral levels (Pisoni & Tash, 1974; Blumstein, Myers, & Rissman, 2005). Phonetic detail provides coarticulatory information that can help listeners identify upcoming words, and data have suggested that listeners use this coarticulatory information on-line in lexical recognition tasks (Gow, 2001). Though one could contend that listeners' ultimate goal is to categorize speech sounds into discrete phonemes, they seem to attend to phonetic detail in the speech signal as well.

The model brings three different analyses of categorical effects together under a single framework. The first piece of this model relates to Huttenlocher et al. (2000)'s account of category effects on visual stimulus reproduction. In their model, when category structure was present in the stimuli, subjects used this structure to compensate for uncertainty in memory traces. We argue that speech perception involves solving the same computational problem as these visual tasks. Parallel to inferring a stimulus value while correcting for memory uncertainty, listeners must infer the phonetic detail of a speaker's target production while correcting for uncertainty in the speech signal.

In addition to drawing parallels between computational problems in speech perception and other areas of cognition, this Bayesian model synthesizes two opposing explanations for the perceptual magnet effect. One account involves speech sound prototypes that act as perceptual magnets, pulling the perception of speech sounds toward them (Kuhl et al., 1992). The idea of a perceptual magnet is formalized in Equation 4, where speech sounds are perceived based on the mean of the category that produced them. The second account ties the perception of speech sounds to the task of inferring category membership (Lacerda, 1995). In line with this, the Bayesian solution to the problem of speech perception with multiple categories (Equation 7) necessitates that listeners first infer category membership. However, in contrast to Lacerda (1995)'s model, which assumes that listeners are perceiving only category membership, the present model predicts that listeners perceive speech sounds in terms of speakers' intended target productions, a continuous variable that depends only partly on category membership. The Bayesian model presented in this paper therefore synthesizes these two previous proposals into a single framework in which the perceptual magnet effect arises through the interaction between shrinkage of perceptual space toward category centers and enhanced discrimination between categories through optimal inference of category membership.

The results presented in this paper establish that the perceptual magnet effect can be explained as the consequence of optimally solving the statistical problem of speech perception using knowledge about the structure of phonetic categories. We are currently conducting empirical tests of the predictions of this model under various parameter manipulations, aiming to differentiate between it and competing models. In addition to providing a basic account of the perceptual magnet effect, this model might be able to shed light on parallels between vowel and consonant perception. It has been noted that within-category discrimination of vowels is easier than that of consonants, and this model suggests that these differences might be related to a higher noise component in consonant perception or a higher amount of meaningful variability in vowel categories. Finally, since the basic assumptions behind the model are not specific to the structure of speech, our Bayesian approach may also provide the foundation for understanding effects of categories on perception more generally, a possibility that we hope to explore in future work.

Acknowledgments We thank James Morgan, Sheila Blumstein, Stefan Benus, Erin Conwell, Sharon Goldwater, Melanie Soderstrom, Elena Tenenbaum, Katherine White, and four anonymous reviewers for helpful comments and discussion. This work was partially supported by a Brown University Fellowship, NSF-IGERT grant 9870676, NSF grant 0631518, and NIH grant HD32005.

References

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

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.

Damper, R. I., & Harnad, S. R. (2000). Neural network models of categorical perception. *Perception and Psychophysics*, *62*(4), 843-867.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature*, *398*, 203-204.

Diesch, E., Iverson, P., Kettermann, A., & Siebert, C. (1999). Measuring the perceptual magnet effect in the perception of /i/ by German listeners. *Psychological Research*, *62*, 1-19.

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.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110-125.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. New York: Chapman & Hall.

Goldstone, R. L., Lippa, Y., & Shiffrin, R. M. (2001). Altering object representations through category learning. *Cognition*, *78*, 27-43.

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

Guenther, F. H., & Gjaja, M. N. (1996). The perceptual magnet effect as an emergent property of neural map formation. *Journal of the Acoustical Society of America*, *100*(2), 1111-1121.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology*, *129*(2), 220-241.

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. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *Journal of the Acoustical Society of America*, *99*(2), 1130-1140.

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.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., et al. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*, B47-B57.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606-608.

Lacerda, F. (1995). The perceptual-magnet effect: An emergent consequence of exemplar-based phonetic memory. In K. Elenius & P. Branderud (Eds.), *Proceedings of the XIIIth international congress of phonetic sciences* (Vol. 2, p. 140-147). Stockholm: KTH and Stockholm University.

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*(4), 358-368.

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.* San Francisco: W. H. Freeman.

Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception and Psychophysics*, *15*(2), 285-290.

Thyer, N., Hickson, L., & Dodd, B. (2000). The perceptual magnet effect in Australian English vowels. *Perception and Psychophysics*, *62*(1), 1-20.