GENERALIZING A NEUROPSYCHOLOGICAL MODEL OF VISUAL CATEGORIZATION TO AUDITORY CATEGORIZATION OF VOWELS

W. Todd Maddox, Randy L. Diehl, Michelle R. Molis
University of Texas, Austin, TX, U.S.A.

ABSTRACT

Twelve male listeners categorized 54 synthetic vowel stimuli that varied orthogonally in F2 and F3 on a BARK scale into the American English vowel categories /ɪ/, /ʊ/, and /□□/. A neuropsychological model of visual categorization, called the Striatal Pattern Classifier (SPC; [1]) is generalized to the auditory domain, and applied separately to the data from each observer. Performance of the SPC is compared with the successful Normal A Posteriori Probability model (NAPP; [2], [3]) of auditory categorization. Versions of the SPC and NAPP that assume linear response region partitions provided similar accounts of the data. Nonlinear versions of both models provided only small improvements in fit.

1. INTRODUCTION

Is that sound a gunshot or a car backfiring? Was that a playful yell or a yell of distress? Everyday we assign objects and events to distinct categories, and often a correct categorization decision might mean the difference between life and death. Categorization is of fundamental importance to the survival of all organisms.

Much research has been devoted to understanding the perceptual and cognitive operations involved in categorization. The current thinking is that there are at least two categorization systems, and that each is associated with separate brain systems (e.g., [4]). There is general agreement that one system is explicit (i.e., rule- or theory-based), is available to conscious awareness, and likely involves frontal brain structures. There is less agreement about the nature of the implicit system, but a growing body of work suggests that this system involves a gradual strengthening of stimulus-response associations within the striatum (a region of the basal ganglia that contains the caudate nucleus and the putamen), and is not available to conscious awareness (e.g., [4]). For example, striatal-damaged Parkinson's disease (PD) patients demonstrate impaired probabilistic (implicit) classification learning, fMRI studies with normal individuals suggest that the striatum is activated during the same task, and animal studies implicate the striatum in category learning ([5] - [8]).

Ashby and Waldron ([1]) recently proposed a neuropsychological model of the implicit categorization system with visual stimuli. Briefly, it is assumed that the stimuli are represented perceptually in higher-level visual areas, perhaps inferotemporal cortex (IT). Since as many as 10,000 cells in visual cortex project to the same striatal cell ([9]), it is assumed that a low-resolution map of the perceptual space is represented among the striatal units. As the observer gains experience with the task each striatal unit becomes associated with a particular response. Thus, the striatum can be thought of as associating a response with a cluster of visual cortical cells. Ashby and Waldron ([1]) referred to this as the striatal pattern classifier (SPC).

The goals of the current study are several. First, we examine the neuropsychological plausibility of the SPC as applied to auditory categorization since most real-world auditory categorization problems are likely solved implicitly. Second, we apply the model to data from a vowel categorization study. Third, we compare the SPC with the highly successful Normal A Posteriori Probability (NAPP; [2], [3]) model of auditory classification. Our goal is not to compare the SPC to NAPP in order to reject one model or the other, but rather to use the NAPP model as a benchmark against which to compare the SPC. If the SPC provides a reasonable account of the data relative to the NAPP model, then we will deem the SPC worthy of additional study. By applying the SPC to auditory categorization, we are taking an important step toward bridging the gap between studies of visual and auditory categorization (see also [10]), and toward offering a general neuro-biological framework for studying both visual and auditory categorization.

1.1 Striatal Pattern Classifier (SPC)

The SPC can be viewed as a neuropsychological implementation of the decision process assumptions of Ashby and Townsend's ([11]) General Recognition Theory (GRT; also called decision bound theory; [12]). GRT is a generalization of signal detection theory to stimuli that vary along multiple dimensions. In GRT, perceptual and decisional processes are characterized by separate and uniquely identifiable parameters.

1.1.1 Perceptual Processes

Real-world objects vary along multiple continuous-valued dimensions such as size, color, fundamental frequency, and amplitude. GRT takes as its fundamental axiom that repeated presentations of the same stimulus yield different perceptual effects (i.e., perceptual noise exists), and assumes that a single multidimensional stimulus can be represented perceptually by a multivariate probability distribution ([11]). For a two-dimensional stimulus, a bivariate normal distribution is assumed to describe the set of percepts. A bivariate normal distribution is described by a mean and variance along each dimension, as well as a covariance term, μ_x, μ_y, σ^2_x, σ^2_y, cov_{xy}, where the subscripts x and y denote dimensions x and y. Figure 1a depicts hypothetical equal likelihood contours for nine stimuli constructed from the factorial combination of three levels along two dimensions x and y. With bivariate normal distributions the equal likelihood contours are always circular or elliptical. The spread of the contour along the x and y dimensions is related to the perceptual variability along each dimension. Notice that the three contours associated with level 1 of dimension y have less variability than the three contours associated with level 2 of dimension y. *Perceptual independence* holds for a single stimulus if and only if the perceptual effects for dimensions x and y are statistically independent (e.g., [11], [13]). With bivariate normal distributions, perceptual independence holds when the major and minor axes of the contour are parallel to the coordinate axes (i.e., when the covariance is zero). Perceptual independence holds for all but the top left, and top right contours in Figure 1a. A positive slope for the major axis implies a positive perceptual dependence (as seen for the top right stimulus), and a negative slope implies a negative perceptual dependence (as seen for the top left stimulus)[1].

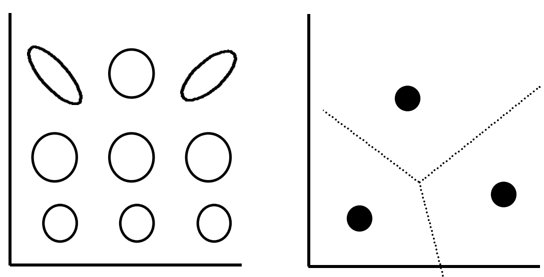

Figure 1. (a) Hypothetical contours of equal likelihood for nine stimuli constructed from the factorial combination of three levels along two stimulus components. (b) Hypothetical response regions and striatal units for three categories.

1.1.2 Decision Processes

In GRT, the experienced observer learns to divide the perceptual space into response regions, and assigns a response to each region. The partitions between response regions are called *decision bounds*. On each trial the observer determines the location of the perceptual effect and gives the response associated with that region of the perceptual space. The SPC offers a neuropsychological theory of this process as applied to the visual categorization. Stimuli are represented perceptually in higher-level visual areas (e.g., IT) and project, in a many-to-one fashion, to striatal cells ([9]). This yields a low-resolution map of the perceptual space among the striatal units. Over trials, each striatal unit becomes associated with a particular categorization response. Thus, the striatum can be thought of as associating a categorization response with a cluster of visual cortical cells. Hypothetical response regions and striatal units for a three-category problem for the nine stimuli in Figure 1a are displayed in Figure 1b. In GRT, the perceptual representation parameters (see Figure 1a) are separate from the response region or striatal unit parameters (see Figure 1b).

1.1.3 A Generalization to Auditory Categorization

Although it is straightforward mathematically to apply the SPC to data from an auditory categorization task, it is important to determine whether there is any empirical support for the neuropsychological plausibility of the SPC as a model of auditory categorization. Although less research has been conducted in the auditory domain, at least three recent studies are relevant. Two recent animal studies found direct projections from auditory cortex into the striatum ([15], [16]), suggesting that a low-resolution map of the auditory perceptual space is represented among the striatal units. If fact, there is some evidence that multiple sensory systems converge in the striatum ([15]). A third study implicated the striatum in categorization of auditory stimuli in the rat ([17]). The rats had to categorize an auditory cue (a 1000 or 8000Hz tone, 75dB) that signaled the location of a food reward. The rats successfully learned the task, and significant striatal involvement was found. Thus, there does appear to be neuroanatomical, neurophysiological, and behavioral data implicating the striatum in auditory processing and categorization. In light of this fact, our next step is to examine quantitatively the ability of the SPC to account for data from an auditory categorization task in normal human listeners. Before introducing the vowel categorization experiment, and the modeling results, we briefly review the NAPP model of auditory categorization.

1.2 Normal A Posteriori Probability (NAPP) Model

Briefly, the NAPP model ([2], [3]) assumes that listeners focus on the output of "detectors" or "filters" that are tuned to language-specific vowel categories. These detectors produce normally distributed output that corresponds roughly to the likelihood of the vowel category give the stimulus input. These likelihoods are then combined using the Luce choice rule (also called the relative goodness rule) to generate predicted response probabilities for each vowel category. From the output of the choice rule one can generate a response surface, and partitions that separate the response regions. These behave much like the decision bounds in the SPC.

1.3 Vowel Categorization Experiment

Listeners labeled a set of synthetic vowel stimuli that varied orthogonally in F2 and F3 and that ranged perceptually among the categories /ɪ/, /ʊ/, and /□□/.

[1] A second form of perceptual interaction in GRT is *perceptual separability* that holds when the perceptual effects of a component are unaffected by the level of the other component (e.g., [14]). In Figure 1a, dimension y is perceptually separable from dimension x, but dimension x is not perceptually separable from dimension y. To provide a straightforward comparison with the NAPP model, all SPC models tested assumed perceptual separability along both dimensions.

2. METHODS

2.1 Listeners

Listeners were twelve males (age range: 18-37; M = 27.2). An attempt was made to limit listeners to a relatively homogeneous dialect group (Central Texas); all listeners were raised in or around the metropolitan areas of Austin, Houston, or Dallas, Texas. They received monetary compensation for their participation.

2.2 Stimuli

Fifty-four, five-formant synthetic vowel stimuli were synthesized using a KLATT88-type cascade resonance synthesizer implemented on a PC. The stimulus space encompassed the American English vowel categories /ɪ/, /ʊ/, and /□□/. Stimuli shared a common first formant (F1) frequency, but varied orthogonally in second and third formant frequency (F2 and F3) on a perceptually motivated frequency scale (Bark). F2 varied between 9.0 and 13.8 Bark (1081 Hz and 2390 Hz) in equal 0.4 Bark steps. F3 varied from 10.0 to15.2 Bark (1268 Hz to 2783 Hz), also in equal 0.4 Bark steps. The Hz values used as input parameters to the synthesizer were calculated to correspond with their desired Bark value equivalents using an equation from Traunmüller ([18]). A plot of the 54 stimuli in the F2, F3 space is displayed in Figure 2. The frequency values of F1, F4, and F5 were 4.5 Bark (455 Hz), 16.2 Bark (3250 Hz) and 17.0 Bark (3700 Hz) respectively. The default bandwidths of the synthesizer were used: 60 Hz for F1, 90 Hz for F2, 150 Hz for F3 and 200 for F4 and F5. All stimuli were 225 ms in length. This duration was chosen to be similar to the measured intrinsic duration of these vowels produced in citation form ([19]). The fundamental frequency was a constant 132 Hz for the initial 150 ms and thereafter fell linearly to 127 Hz over the final 75 ms. Stimuli were ramped on and off with a 10 ms half-cosine function and were normalized for RMS amplitude.

2.3 Procedure

Listeners were seated at separate response stations in a sound attenuated chamber. Stimuli were presented over Beyer DT-100 headphones at a level of 70 dB SPL. Fourteen randomized blocks of the 54 stimuli were presented (756 trials/listener). Listeners identified each stimulus by pressing one of three response buttons labeled with the key words "hid", "hood", and "heard" corresponding to the three identifiable American English vowel categories.

3. RESULTS AND THEORETICAL ANALYSES

3.1 Accuracy Analyses

The proportions of /ɪ/, /ʊ/, and /□□/ responses to each stimulus were computed for each listener and used for the model-based analyses. Figure 2b displays the most common response for each of the 54 stimuli averaged across observers.

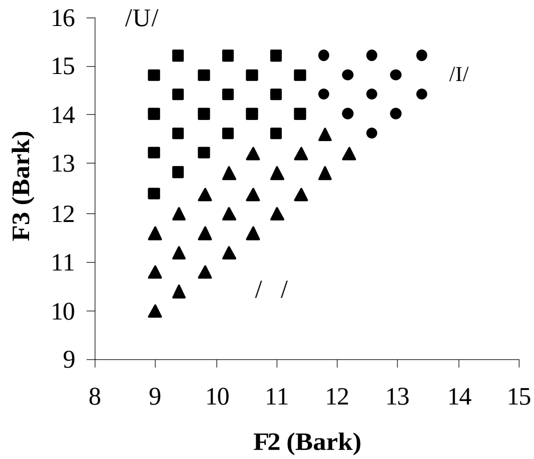

Figure 2. Schematic illustration of the 54 synthetic vowel stimuli in the F2, F3 space, and symbolic representation of the most common vowel category response (/ɪ/, /ʊ/, and /□□/) for each of the 54 stimuli averaged across observers. Circle = /ɪ/, square = /ʊ/, and triangle = /□□/.

3.2 Model-based Analyses

The data to be modeled were the observed vowel categorization response frequencies (for categories “hid”, “hood”, and “heard”) for each of the 54 stimuli, resulting in 108 degrees of freedom in the data for each listener. All model-based analyses were performed at the level of the individual listener. Data were not collapsed across listeners since averaging often alters the structure of the categorization data in such a way that the correct model of individual performance provides a poor account of the aggregate data ([20]).

The model parameters (to be described shortly) were estimated using maximum likelihood procedures ([21]) and the goodness-of-fit statistic was

$$AIC = 2r - 2lnL,$$

where r is the number of free parameters and L is the likelihood of the model given the data (e.g., [21]). The AIC statistic penalizes a model for extra free parameters in such a way that the smaller the AIC, the closer a model is to the “true model,” regardless of the number of free parameters. Thus, to find the best model among a given set of competitors, one simply computes an AIC value for each model, and chooses the model associated with the smallest AIC value.

3.2.1 Striatal Pattern Classifier

Perceptual Representation Assumptions: Three sets of perceptual representation assumptions were tested. All models tested made the following two assumptions:

(1) The mean perceptual effects (i.e., the mean for each bivariate normal distribution of perceptual effects) were located at the synthesized F2, F3 values (see Figure 2a).

(2) The perceptual covariance matrices were identical across stimuli. In other words, the F2 perceptual variability was constant across the 54 stimuli, the F3 perceptual variability was constant across the 54 stimuli, and the F2, F3 perceptual dependence was constant across the 54 stimuli. This is referred to as a Stimulus Invariant (SI) perceptual representation since the perceptual covariance matrix entries are invariant across stimuli.

The three perceptual representations tested differed only in the assumptions made about the F2 and F3 perceptual variabilities, and the F2, F3 perceptual dependence. In addition, the three models were "nested" in the sense that a more restricted model could be derived from a more general model by setting some of the parameters of the more general model to constants. The perceptual representation assumptions from most restricted to most general are outlined as follows:

<u>SPC(σ, σ, 0)</u>: This model assumes that the perceptual variance along the F2 dimension is equivalent to the perceptual variability along the F3 dimension, and that perceptual independence is satisfied between the perceived F2, F3 values (1 parameter). In other words, the contours of equal likelihood for all 54 stimuli are circles of the same diameter.

<u>SPC(σ_{F2}, σ_{F3}, 0)</u>: This model assumes that the perceptual variance along the F2 dimension is different from the perceptual variability along the F3 dimension, and that perceptual independence is satisfied between the perceived F2, F3 values (2 parameters). In other words, the contours of equal likelihood for all 54 stimuli are ellipsoids of the same size and orientation whose major and minor axes are parallel to the F2, F3 axes.

<u>SPC(σ_{F2}, σ_{F3}, $\rho_{F2,F3}$)</u>: This model assumes that the perceptual variance along the F2 dimension is different from the perceptual variability along the F3 dimension, and that there is a perceptual dependence between the perceived F2, and F3 values (3 parameters). In other words, the contours of equal likelihood for all 54 stimuli are ellipsoids of the same size and orientation whose major and minor axes are NOT required to be parallel to the F2, F3 axes.

Response Region Assumptions: For simplicity, we assumed that a single striatal unit was associated with each of the three vowel categories, for a total of three striatal units. Under these conditions, the decision bounds that partition the F2, F3 space into three vowel categories are linear. The location of one striatal unit in the F2, F3 space can be fixed a priori, yielding a total of four free striatal unit parameters. We also applied a version of the SPC that assumed two striatal units per category (10 free parameters). This yields piece-wise linear decision bounds, and will be compared with the non-linear version of the NAPP model. This model was fit only under the most general perceptual representation assumptions [i.e., SPC(σ_{F2}, σ_{F3}, $\rho_{F2,F3}$)]. We also fit a version of the SPC that allowed quadratic response region boundaries. This model did not provide a significant improvement in fit over the piece-wise linear SPC for any observer and will not be discussed further.

An iterative search routine that minimized the AIC value was used to estimate the model parameters (see [40] for details).

3.2.2 NAAP Model

It is straightforward to show that the linear version of the NAPP model is equivalent to a system of three linear logistic equations ([3]), and that the quadratic version of the NAPP model is equivalent to non-linear logistic models. Thus, both can be fit using readily available "canned" statistical software.

The number of free parameters, AIC, and percent of variance accounted for by the three linear SPC, the piece-wise linear SPC, linear NAPP, and the non-linear NAPP models averaged across observers are presented in Table 1. We begin by focusing on the linear versions of the SPC and NAPP models. The results can be summarized as follows. First, the fits of the SPC are comparable to those of the NAPP. Although not shown in the Table, for six of the 12 observers the SPC provided a superior account of the data (based on AIC). This finding lends support to the SPC as a reasonable model of vowel categorization. Second, restricting attention to the SPC models, the SPC(σ_{f2}, σ_{f3}, 0) provided an improvement in fit over the simpler SPC(σ, σ, 0) for only 2 of 12 observers, whereas the SPC(σ_{f2}, σ_{f3}, $\rho_{f2,f3}$) provided a superior fit for 9 of 12 observers. Based on a sign test, this latter difference is significant ($p < .05$). Finally, notice that the fit of the SPC(σ_{f2}, σ_{f3}, $\rho_{f2,f3}$) and the NAPP models are similar and quite good accounting for 96.10% and 95.59% of the variance in the observed data, respectively.

Table 1. Goodness-of-Fit (AIC) and Percent of Variance Accounted for (R^2) Averaged Across Observers for the SPC and NAPP Models.

Model	# parms	AIC	R^2
Linear-SPC(σ,σ,0)	5	187.29	95.24
Linear-SPC(σ_{F2}, σ_{F3}, 0)	6	185.48	95.58
Linear-SPC(σ_{F2}, σ_{F3}, $\rho_{F2,F3}$)	7	175.89	96.10
Linear-NAPP	6	171.19	95.59
Nonlinear-SPC(σ_{F2}, σ_{F3}, $\rho_{F2,F3}$)	13	175.66	96.75
Nonlinear-NAPP	12	151.55	97.68

To gain a more detailed understanding of the SPC(σ_{f2}, σ_{f3}, $\rho_{f2,f3}$), we examined the perceptual representation and response region parameters. There were moderate individual differences in the resulting perceptual representation parameters, but two general statements can be made. First, the perceptual variability along the F3 dimension was generally larger than the perceptual variability along the F2 dimension ($p = .073$), suggesting that the auditory system in the current task was more sensitive to the F2 dimension. Second, in general there was a negative correlation between the perceived F2 and F3 values ($p = .073$), suggesting that high perceived values of F2 were generally associated with low perceived values of F3. A visual examination of the best fitting response regions for the SPC(σ_{f2}, σ_{f3}, $\rho_{f2,\ f3}$), and NAPP models suggested that the response regions were quite similar for the two models.

The nonlinear NAPP model provided a better account of the data than the linear NAPP model (based on AIC), but the improvement in percent of variance accounted for was small (linear NAPP: 95.59%; nonlinear NAPP: 97.68%). The nonlinear (piece-wise linear) SPC provided essentially no improvement in fit based on AIC, and only a small improvement based on percent of variance accounted for (linear SPC: 96.10%; piece-wise linear SPC: 96.75%). Taken together, these results suggest that additional flexibility afforded by nonlinear decision rules was not necessary to capture the listener's vowel categorization behavior.

4. GENERAL DISCUSSION

This article reports the results of an auditory vowel categorization experiment in which listeners classified 54 synthetic vowel stimuli that varied along the F2 and F3 dimensions into one of three vowel categories /I/, /U/, and /□□/. A successful, neuropsychological model of categorization in the visual domain, the SPC, was generalized to the auditory domain, and was applied separately to each listener's data from the auditory vowel categorization task. The SPC was compared with the benchmark NAPP model. Both models provided good and roughly equivalent accounts of the data, suggesting that the SPC does generalize to the auditory domain. This finding is important because it suggests that a model with a reasonable neuro-biological architecture can be applied in both the visual and auditory domains.

4.1 SPC Perceptual Representation Parameters

The best fitting parameters from the most parsimonious SPC model suggested that there was more variability in the perceived F3 values than in the perceived F2 values. In addition, the model parameters indicated that high perceived values of F2 were associated with low perceived values of F3—that is, there was a negative perceptual dependence between perceived F2 and F3 values. This latter effect may reflect a type of auditory integration whereby closely spaced formants are averaged into a single spectral prominence with a "center of gravity" located between the nominal values of the two formants ([45]). By this account, raising the frequency of the second formant such that it enters the region of spectral integration with the third formant would, in effect, lower the perceived value of F3. Independent evidence that such spectral integration may be occurring in the present experiment is that the average boundary between the categories /U/, and /□□/ occurs at an F3-F2 distance (nearly 3 Bark) that is close to Chistovich and Lublinskaya's ([22]) proposed bandwidth of spectral integration. Syrdal ([23]) proposed that this critical limit of integration yields a quantal shift in perception ([24]) well suited for the location of vowel category boundaries.

4.2 Linearity of Response Region Boundaries

One interesting finding from the current study is that listener's performance was well captured by SPC and NAPP models that assumed linear partitions between the three vowel categories. In most categorization situations, the optimal classifier (i.e., the hypothetical device that maximizes long-run accuracy) uses nonlinear partitions of the response regions. Yet in the current study, linear partitions provided an excellent account of the listener's data. The use of linear as opposed to nonlinear partitions is not without precedence. For example, in the visual domain, Ashby, et al ([25]) found strong evidence that observers used linear partitions to solve a categorization problem in which nonlinear partitions were optimal. Boundary linearity has also been reported in several earlier studies of Swedish ([26]), Russian ([27]), and German ([28]) vowel perception.

4.3 Relations Between the SPC and NAPP models

Although the SPC and NAPP models provided similar accounts of the current data, the architecture of the two models is quite different. One important distinction between the SPC and the NAPP models is that the SPC assumes that repeated presentations of a single stimulus yield different perceptual effects (i.e., perceptual noise exists). The NAPP model makes no strong claim about the perceptual processing of individual stimuli, although in its present form it assumes implicitly no perceptual noise. Thus, the SPC assumes a probabilistic perceptual representation whereas the NAPP model makes deterministic perceptual representation assumptions.

The SPC assumes that the observer determines the location of the perceptual effect on each trial and gives the categorization response associated with that region of the perceptual space. Thus, each location (or perceptual effect) in the space is deterministically associated with one of the categorization responses. The NAPP model, on the other hand, assumes that each stimulus is processed by a set of "detectors" or "filters" that are tuned to language-specific vowel categories. These detectors produce normally distributed outputs that correspond roughly to the likelihood of the vowel category give the stimulus input. These category likelihoods are then combined using the relative goodness rule to determine a categorization response probability for each category. Thus, each location in the space is probabilistically associated with one of the categorization responses. Further work is needed to determine whether a probabilistic perceptual representation along with deterministic decision processes (as in the SPC), or a deterministic perceptual representation along with probabilistic decision processes (as in the NAPP) is more parsimonious with respect to the neuro-biological and behavioral data.

5. ACKNOWLEDGEMENTS

This research was supported by NSF Grant SBR-9796206, and NIMH Grant R01 MH59196 to WTM, and by NIDCD Grant DC00427-11 to RLD. Corresponding author: W. Todd Maddox, Department of Psychology, Mail Code B3800, University of Texas, Austin, Texas, 78712. E-mail: maddox@psy.utexas.edu.

6. REFERENCES

[1] Ashby, F.G., & Waldron, E.M. (1999). On the fundamental nature of human pattern classification. *Psychonomic Bulletin and Review*, *6*, 363-378.

[2] Nearey, T.M. (1990). The segment as a unit of speech perception. *Journal of Phonetics*, *18*, 347-373.

[3] Nearey, T.M., & Hogan, J. (1986). Phonological contrast in experimental phonetics: Relating distributions of production data to perceptual categorization curves. In *Experimental phonology*. J. Ohala, J. Jaeger (Ed.), pp. 141-161, New York: Academic Press.

[4] Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442-481.

[5] Knowlton, B.J., Mangels, J.A., & Squire, L.R. (1996). A neostriatal habit learning system in humans. *Science*, *273*, 245-254.

[6] Maddox, W.T., & Filoteo, J.V. (in press). Striatal contributions to category learning: Quantitative modeling of simple linear and complex non-linear rule learning in patients with Parkinson's Disease. *Journal of the International Neuropsychological Society*.

[7] Poldrack, R.A., Prabhakaran, Seger, C.A., & Gabrieli, J.D.E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology*, *13*, 564-574.

[8] McDonald, R.J., & White, N.M. (1993). A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience*, *107*, 3-22.

[9] Wickens, J. (1993). *A theory of the striatum.* New York: Pergamon Press.

[10] Kingston, J. & Macmillan, N.A. (1995). Integrality of nasalization and F_1 in vowels in isolation and before oral and nasal consonants: A detection-theoretic application of the Garner paradigm. *Journal of the Acoustical Society of America*, *97*, 1261-1285.

[11] Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154-179.

[12]Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception and Psychophysics*, *53(1)*, 49-70.

[13] Ashby, F. G., & Maddox, W. T. (1991). A response time theory of perceptual independence. In J.P. Doignon & J.C. Falmagne (Eds.). *Mathematical psychology: Current developments*. (pp. 389-414). Springer Verlag.

[14] Maddox, W.T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional Models of Perception and Cognition*. (pgs. 147-180). Erlbaum: Hillsdale, NJ.

[15] Chudler, E.H., Sugiyama, K., & Dong, W.K (1995). Multisensory convergence and integration in the neostriatum and globus pallidus of the rat. *Brain Research*, *674*, 33-45.

[16] Arnauld, E., Jeantet, Y., Arsaut, J., Desmotes-Mainard, J. (1996). Involvement of the caudal striatum in auditory processing: *c-fos* response to cortical application of picrotoxin and to auditory stimulation. *Molecular Brain Research*, *41*, 27-35.

[17] Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, & Graybiel, A.M. (1999). Building neural representations of habits. *Science*, *286*, 1745-1749.

[18] Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America*, *88*, 97-100.

[19] Hillenbrand, J.M., Getty, L.A., Clark, M.J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099-3111.

[20] Maddox, W.T. (1999). On the danger of averaging across observers when comparing decision bound and generalized context models of categorization. *Perception & Psychophysics*, *61*, 354-374.

[21] Ashby, F. G. (1992). Multivariate probability distributions. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 1-34). Hillsdale, NJ: Erlbaum.

[22] Chistovich, L.A., & Lublinskaya, V.V. (1979). The "center of gravity" effect in vowel spectra and critical distance between the formants. *Hearing Research 1*, 185-195.

[23] Syrdal, A.K. (1985). Aspects of a model of the auditory representation of American English vowels. *Speech Communication*, *4*, 121-135.

[24] Stevens, K.N. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*, 3-45.

[25] Ashby, F.G., Waldron, E.M., Lee, W.W., & Berkman, A. (in press). Suboptimality in categorization and identification, *Journal of Experimental Psychology: General*.

[26] Carlson, R., Granström, B., & Fant, G. (1970). Some studies concerning perception of isolated vowels. *Speech Transmission Laboratory: Quarterly Progress and Status Report*, *2-3*, 19-35.

[27] Karnickaya, E.G., Mushikov, V.N., Slepokurova, N.A., & Zhukov, S.J. (1975). Auditory processing of steady-state vowels. In G. Fant, & M.A.A. Tathan, (Eds.), *Auditory analysis and the perception of speech* (pp. 37-53). New York: Academic Press.

[28] Hose, B., Langner, G., & Scheich, H. (1983). Linear phoneme boundaries for German synthetic two-formant vowels. *Hearing Research*, *9*, 13-25.