Bayesian Approaches to Perception

Randy L. Diehl and Sarah C. Sullivan
University of Texas at Austin

Rev. Thomas Bayes (1702-1761)

Bayesian Statistical Decision Theory (Ideal Observer Theory)

- Describes how to perform a perceptual or cognitive task optimally, given the available information and any constraints that apply.

Kinds of tasks for which ideal observers have been derived

- Coding
- Detection
- Discrimination
- Estimation
- Object recognition
- Categorization
- Navigation

A simple categorization task

- Assume there are n stimulus categories $c_1, c_2, \ldots, c_n$, and the observer's task is to identify the category, given a particular stimulus $\mathbf{S}$.

- The optimal decision rule is:

If $p\left(c_i \mid \mathbf{S}\right) > p\left(c_j \mid \mathbf{S}\right)$ for all $j \neq i$, then pick c_i

Posterior probabilities

How are these posterior probabilities calculated?

- Bayes' Theorem

Bayes' Theorem

$$p\left(c_i \middle| \mathbf{S}\right) = \frac{p\left(\mathbf{S} \middle| c_i\right) p\left(c_i\right)}{p\left(\mathbf{S}\right)}$$

$p\left(c_i \middle| \mathbf{S}\right)$ = posterior probability

$p\left(\mathbf{S} \middle| c_i\right)$ = likelihood

$p\left(c_i\right)$ = prior probability

$p\left(\mathbf{S}\right)$ = normalizing constant

An Essay Towards Solving a Problem in the Doctrine of Chances.
Philosophical Transactions of the Royal Soc. of London, **53**, pp370-418, 1763.

More generally:

optimal decision rules are expressed in terms of prior probabilities, likelihoods, and *utility functions*, which specify benefits and costs of picking the right or wrong category.

Ideal observer theory provides a rigorous framework for

- representing physical and statistical properties of the environment
- describing the perceiver's task
- deriving computational theories of optimal performance

But perceptual performance of humans and other animals is generally less than optimal.

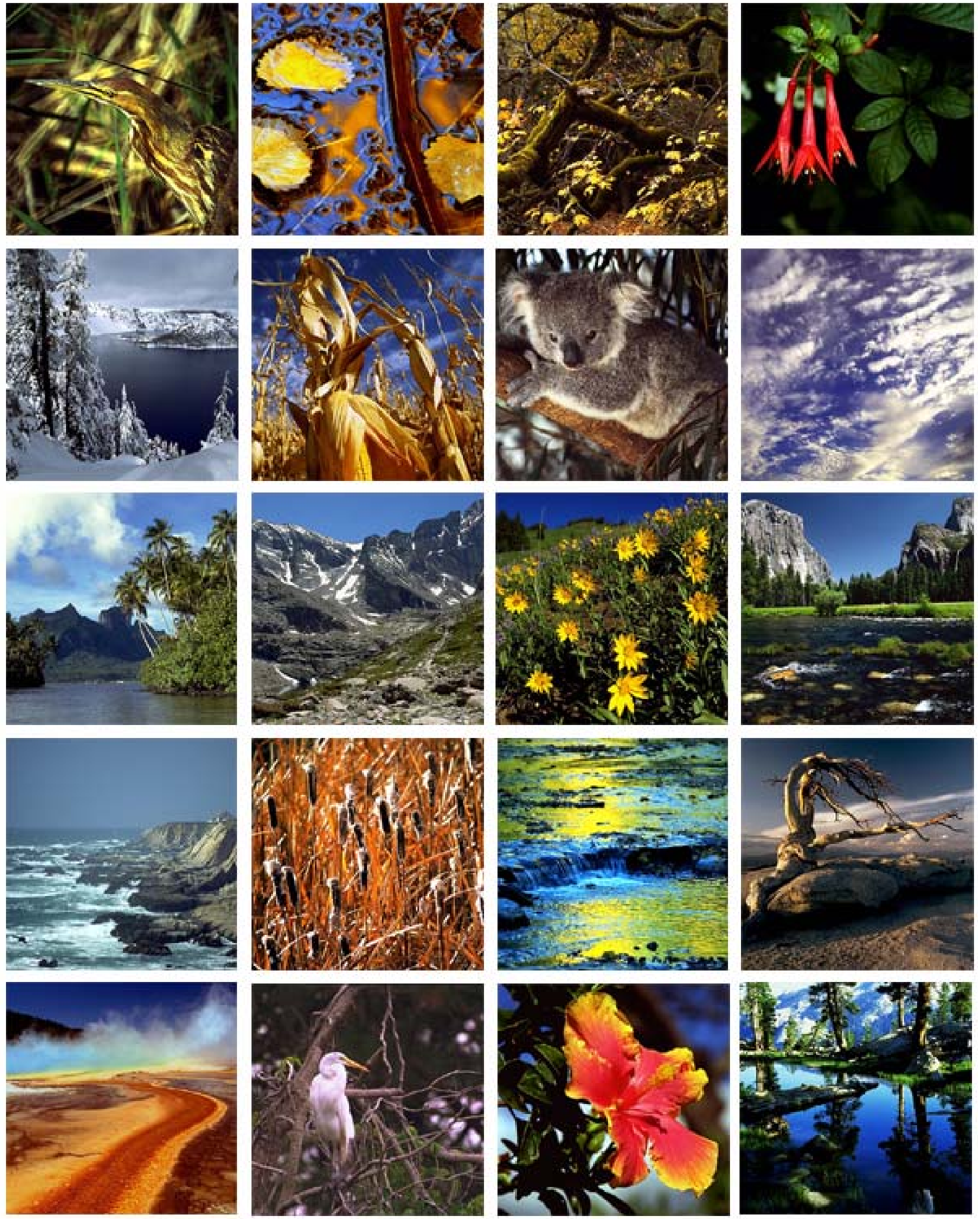
So why bother deriving ideal observers?

Ideal observers provide

- an appropriate benchmark for evaluating actual performance
- a useful starting point for developing models of actual performance

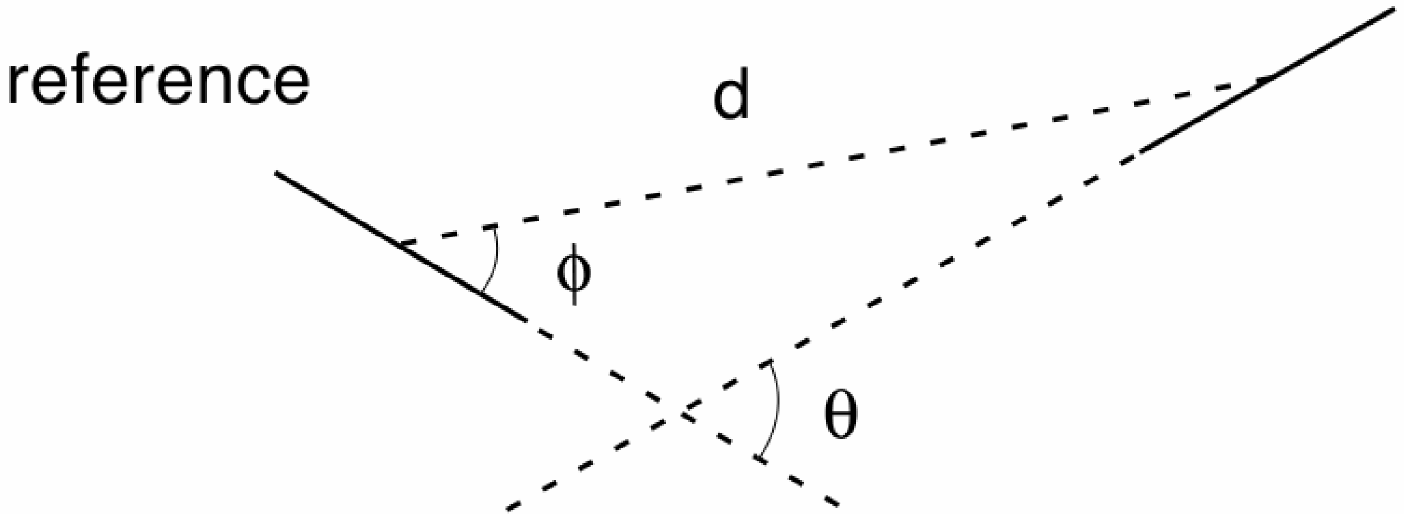
An example from vision

- Edge co-occurrence in natural images predicts contour grouping performance
  (Geisler, Perry, Super, & Gallogly, *Vision Res*, 2001, 41, 711-724)



Extracting edge elements from a scene

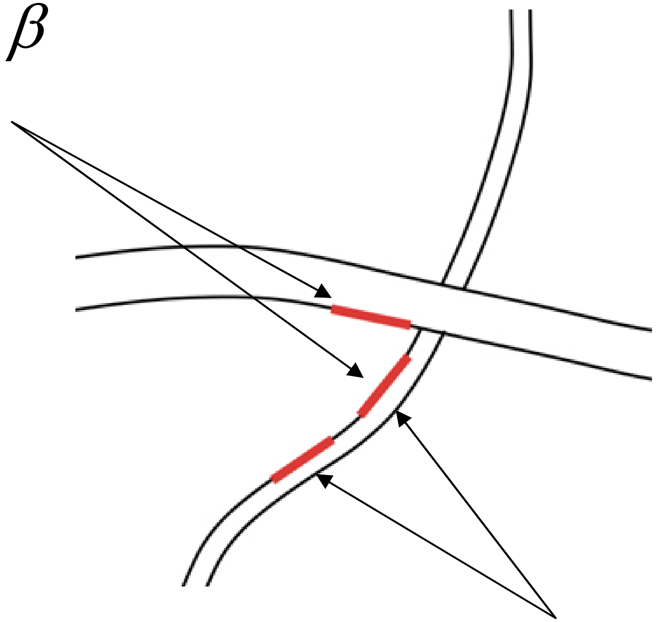
Geometrical relations between two edge elements



Contour grouping

$$\frac{p(\mathbf{S}|c)}{p(\mathbf{S}|\sim c)} < \beta$$

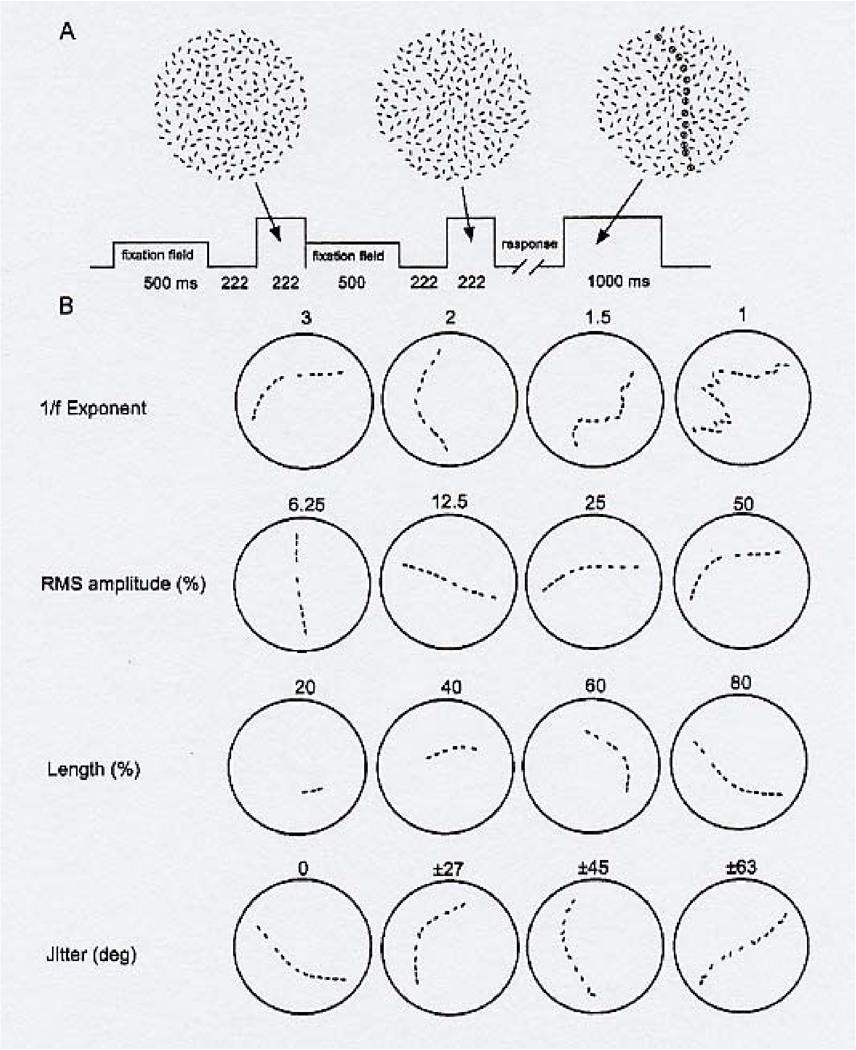
$$\mathbf{S} = (d, \phi, \theta)$$



$$\beta = \frac{p(\sim c)}{p(c)}$$

$$\frac{p(\mathbf{S}|c)}{p(\mathbf{S}|\sim c)} > \beta$$

Contour grouping task



Comparing human and ideal performance

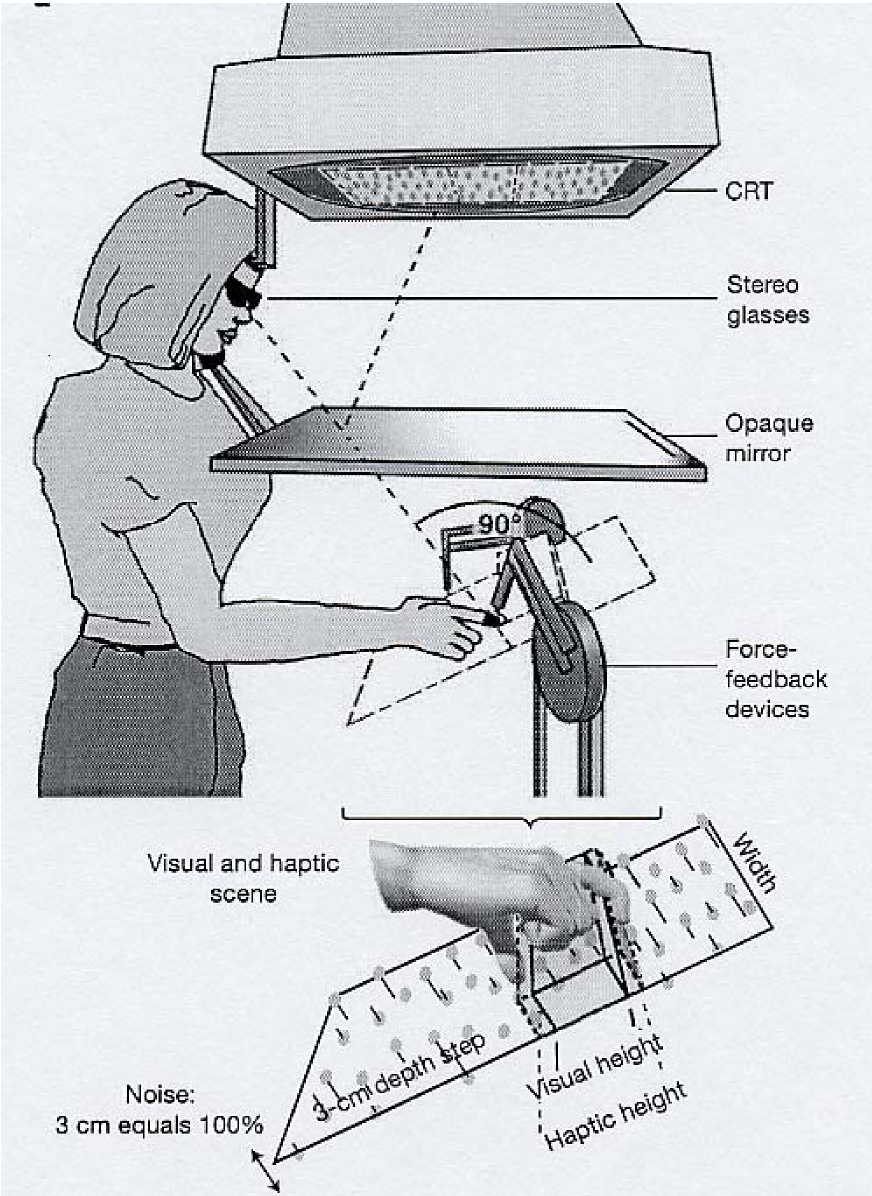
- ***Results:*** Correlation between human and ideal performance = +.9
- ***Implication:*** Humans make nearly optimal use of natural scene statistics in performing the contour grouping task.

A Bayesian framework for natural selection

- Geisler, W.S. & Diehl, R.L. *Philosophical Transactions of the Royal Society of London,.Biological Science*., 2002, 357, 419-448.
- Geisler, W.S. & Diehl, R.L. *Cognitive Science*, 2003, 27, 379-402.

Another example

- Humans integrate visual and haptic information in a statistically optimal fashion (Ernst & Banks, *Nature*, 2002, 415, 429-433).



Task : Estimate height of raised surface

Conditions: Visual only, haptic only, both visual and haptic

Vary the relative informativeness of the two sources of information by varying visual noise

Result: Humans integrate the two sources optimally

A category learning experiment

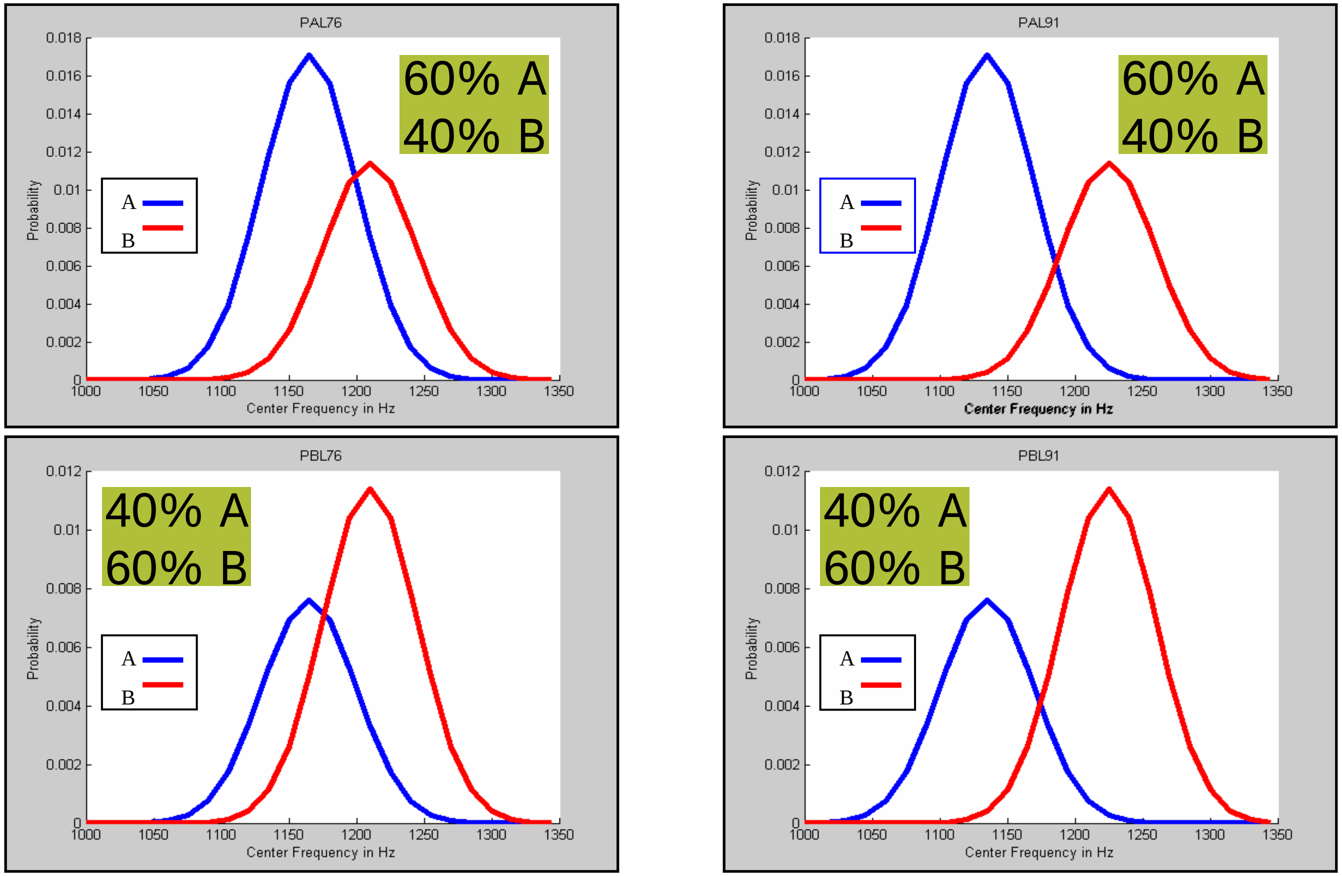
(Sullivan, Lotto, & Diehl)

- Task: Subjects were presented a series of sounds from two overlapping Gaussian-shaped distributions and were asked to identify the sounds as belonging category “A” or “B”. Feedback was provided.

- Stimuli: 25 samples of band-pass white noise with varying center frequencies starting at 1000 Hz, increasing by 15 Hz steps.

- Training: 3 consecutive days of 8 five-minute blocks.

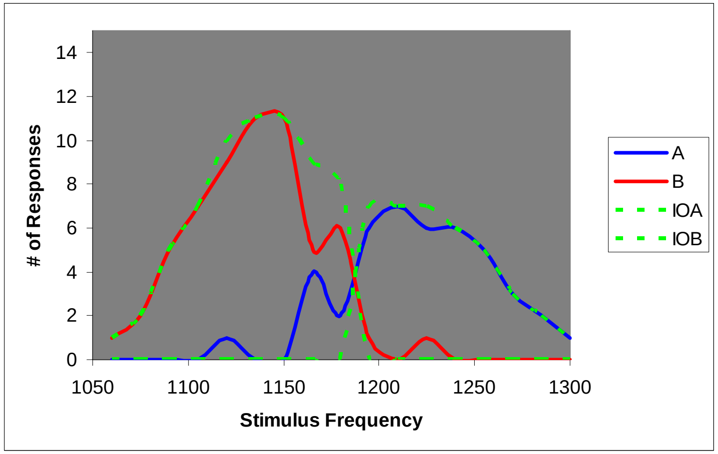
Conditions: Vary priors and likelihood ratios

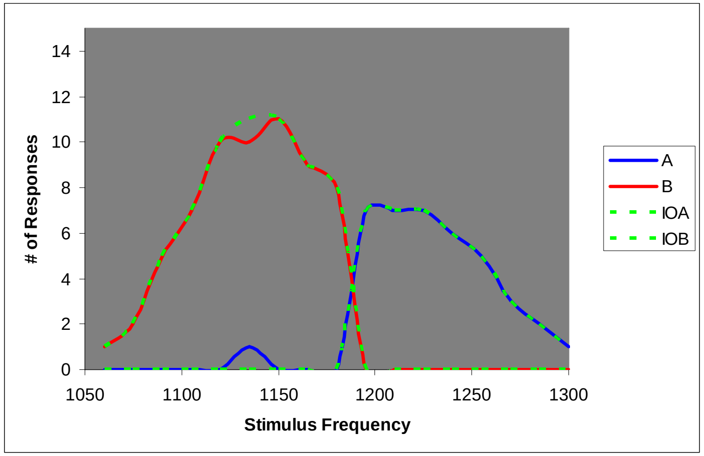


Some preliminary results

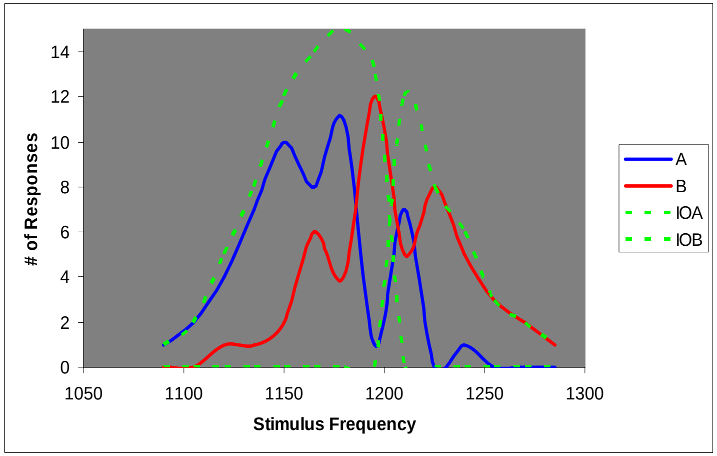
Block 1 **Block 24**

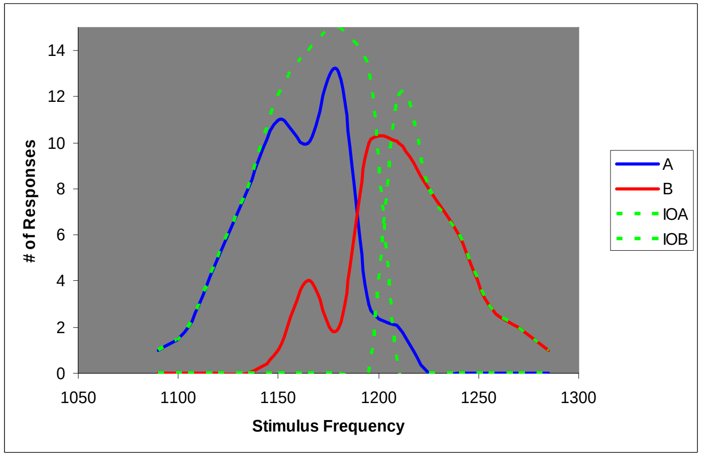
S1





S2





Bayesian modeling of speech perception

- Massaro, Nearey studies
- Measuring the prior probabilities and stimulus likelihoods of phonetic categories in natural speech (cf. Geisler et al)
- Modeling effects of L1 on L2 learning
- Modeling information integration across cues and sense modalities (cf. Ernst and Banks)

Acknowledgments

- Thanks to Andrew Lotto, Lori Holt, Micah Goldwater, Jessica Hay, Momoko Sato, Amy Coren, and Cheryl Moran.

- Work supported by NIDCD grant no. DC00427 to RLD.