Optimal Auditory Categorization on a Single Dimension

Sarah C. Sullivan (sullivans@boystown.org)
Center for Hearing Research, Boys Town National Research Hospital
425 North 30th Street, Omaha, NE, 68131, USA

Andrew J. Lotto (lottoa@boystown.org)
Center for Hearing Research, Boys Town National Research Hospital
425 North 30th Street, Omaha, NE, 68131, USA

Randy L. Diehl (diehl@psy.utexas.edu)
Department of Psychology, University of Texas
1 University Station A8000, Austin, TX, 78712, USA

Introduction

A number of auditory tasks, including speech perception, require listeners to categorize stimuli on the basis of one or more features of the input. In many cases, especially speech, there is no one-to-one mapping between values along continuous features and discrete categories (e.g., phonemes). How then do perceptual systems categorize stimuli under uncertainty? One possible solution is that perceptual systems identify and use statistical information inherent in the acoustic environment. We propose that perceivers incorporate distributional knowledge about the acoustic environment with the information provided by the signal in order to make optimal (i.e., maximized accuracy) categorical decisions. Statistical approaches such as this are widely used in vision research but are rarely applied to auditory or speech perception. Our goal in this study was to develop a framework that will provide testable hypotheses about the nature of statistical (distributional) learning in auditory perception in general and specifically in speech perception.

Methods

In this experiment, participants were presented non-speech sounds sampled from two overlapping distributions. The sounds consisted of 25 narrow-band noise bursts varying in center frequency from 1000 to 1360 Hz. Three different conditions were created by varying parameters of the training distributions. The distributions varied in the ratio of stimuli in each category (i.e., prior probabilities of each category) as well as the amount of overlap between the two distributions. Figure 1 displays the distributions for one of these conditions. The listeners were asked to identify the sounds as belonging to one of two categories ("A" or "B") and feedback was provided after each trial. Due to the overlap between the distributions, there was no deterministic relationship between center frequency and category label for many stimuli.

Results

In decision tasks such as these, optimal performance requires listeners to create a criterion boundary on the dimension (i.e., a particular frequency). Stimuli on either side of this boundary should receive different category labels. Within as few as one block of training trials, most listeners displayed a stable category boundary. Boundaries were estimated from categorization functions averaged across several training blocks. These boundaries varied as a function of the distribution characteristics and were statistically equivalent to the point of distribution intersection. Slopes from categorization functions were steeper than the slopes of the training distributions; suggesting that listeners were more likely using a criterion bound as opposed to simply probability matching the distributions. In general, listeners were responding in a near optimal manner with minimal experience with the training distributions. The ability to use distributional information to map from continuous dimensions to category labels is also essential for speech sound categorization.

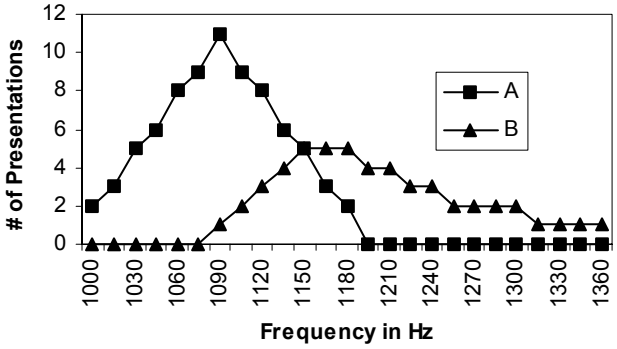

Figure 1: Stimulus Distributions

Acknowledgements

This research was funded by the National Institute on Deafness and Other Communication Disorders and the National Science Foundation.