Perceptual Plasticity in Spatial Auditory Displays

BARBARA G. SHINN-CUNNINGHAM, TIMOTHY STREETER, and JEAN-FRANÇOIS GYSS
Hearing Research Center, Boston University

Often, virtual acoustic environments present cues that are inconsistent with an individual's normal experiences. Through training, however, an individual can at least partially adapt to such inconsistent cues through either short- [Kassem 1998; Shinn-Cunningham 2000; Shinn-Cunningham et al. 1998a, 1998b; Zahorik 2001] or long- [Hofman et al. 1998] term exposure. The type and degree of inconsistency as well as the length of training determine the final accuracy and consistency with which the subject can localize sounds [Shinn-Cunningham 2000]. The current experiments of short-term adaptation measure how localization bias (mean error) and resolution (precision) change when subjects are exposed to auditory cue rearrangements simpler than those previously investigated. These results, combined with those of earlier experiments, suggest that there is plasticity at many different levels of the spatial auditory processing pathway with different time scales governing the plasticity at different levels of the system. This view of spatial auditory plasticity has important implications for the design of spatial auditory displays.

Categories and Subject Descriptors: [**General Literature - General**]—*C fe e ce ceed gs*

General Terms: Experimentation, Human Factors, Performance

Additional Key Words and Phrases: Auditory display, spatial hearing, virtual auditory space

1. INTRODUCTION

A basic goal of spatial auditory display research is to provide listeners with cues that allow accurate localization of sound sources. One approach to achieving this goal is to provide the most realistic cues possible so that the stimuli are essentially identical to the cues heard in a real environment. However, the inherent difficulty and expense of such an approach limits the veridicality that can be achieved in a practical, reasonably priced spatial auditory display. An alternative approach is to try to train listeners to accurately localize even when localization cues are different than normal experience [e.g., see Shinn-Cunningham et al. 1998; Zahorik 2001]. Of course, because the type of discrepancies between normal and altered cues will affect how rapidly and completely subjects can adapt (and ultimately, how well they localize sounds), it is important to explore how these factors are influenced by different cue rearrangements.

2. BACKGROUND

With long-term training, subjects can localize accurately even when the acoustic localization cues are inconsistent with previous experience [e.g., Hofman et al. 1998]. However, short-term training experiments suggest that subjects may be able to rapidly adapt only to linear transformations of auditory space [Shinn-Cunningham 2000; Shinn-Cunningham et al. 1998a, 1998b]. In particular, when subjects

This research was supported by the Whitaker Foundation for Biomedical Research.
Authors' addresses: Barbara G. Shinn-Cunningham or Tim Sreeter, Department of Cognitive and Neural Systems, Boston University, 677 Beacon St., Boston, MA, 02215, USA; http://www.cns.bu.edu/∼shinn; email: {shinn, timstr}@cns.bu.edu.

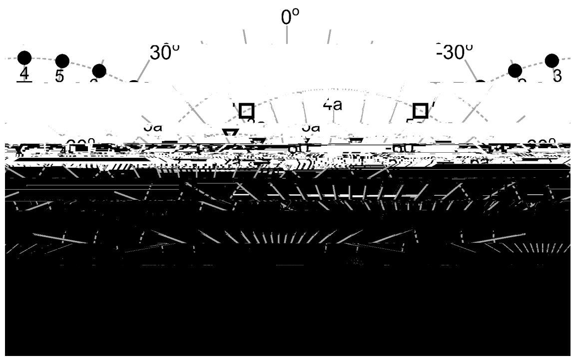

Fig. 1. Nonlinear and linear spatial remappings.

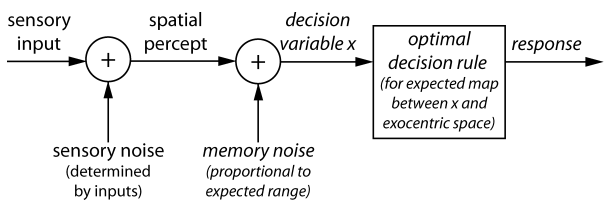

Fig. 2. Block diagram of model of adaptation.

are asked to learn new, nonlinear associations between spatial cues and exocentric locations, residual errors remain even after performance has asymptoted.

Fmgo-zimuth.2. 310n2(HRTFs8()10n)-used78310n2(simulnse8()10n)-the8(310n1jsoucces8(310n2(we-278)10n1jda)30(

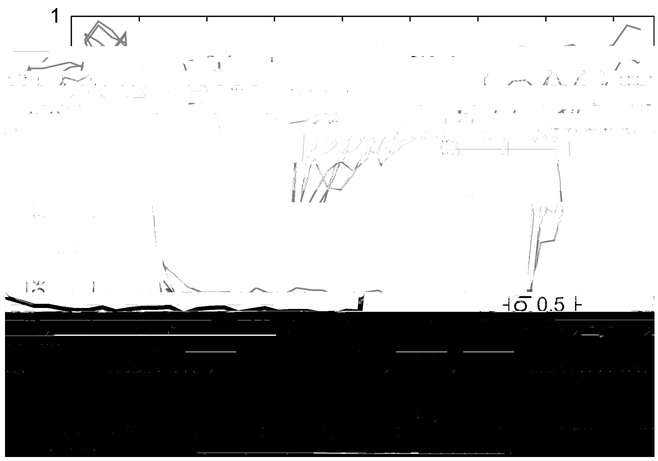

Fig. 3. Slope of mean response versus HRTF position normal cue position. Individual subjects in gray; across subject average in black.

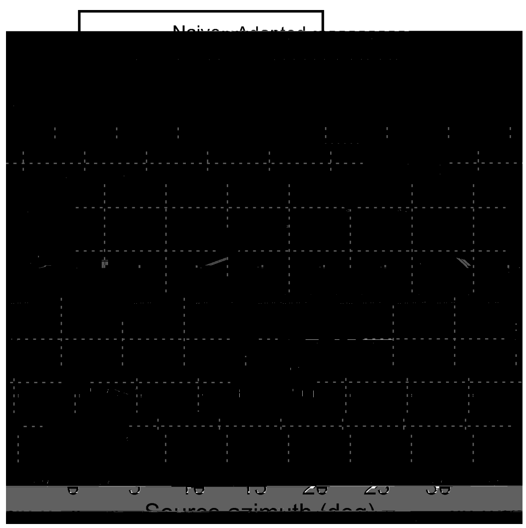

Fig. 4. Bias (mean localization error in units of standard deviation in the response) as a function of source azimuth. Positive bias indicates subject responses err away from the median plane.

experiments, during the altered-cue runs, the slope describing this relationship decreases as subjects adapt to the larger-than-normal cue range. Figure 3 plots these slopes as a function of run number for both individual subjects (gray) and averaged across subjects (black). Similar to earlier nonlinear cue rearrangement experiments [Shinn-Cunningham et al. 1998b], during the altered cue runs, the slope decreases and asymptotes at the “optimal” slope. Unlike previous experiments, in the current experiment, this optimal slope produces a mean square error of zero, since the rearrangement can be perfectly fit by a line.

Figures 4 and 5 show how the bias and resolution (respectively) of subject responses evolves during the experiment. In order to find these values, the raw confusion matrices, which describe the number of times a subject responded that the source was at location i given that the source was actually from

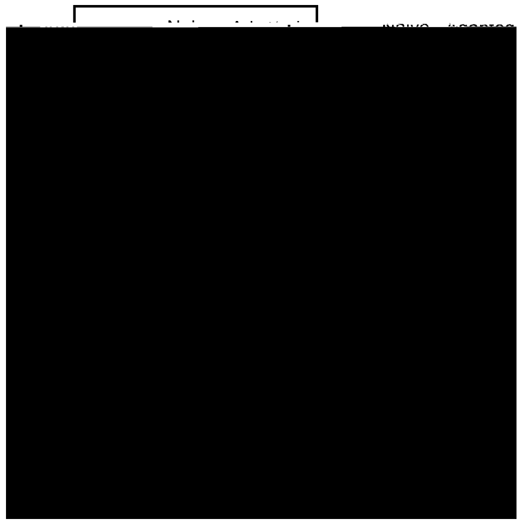

Fig. 5. Resolution (difference in mean responses for adjacent locations in units of standard deviation in responses) as a function of source azimuth.

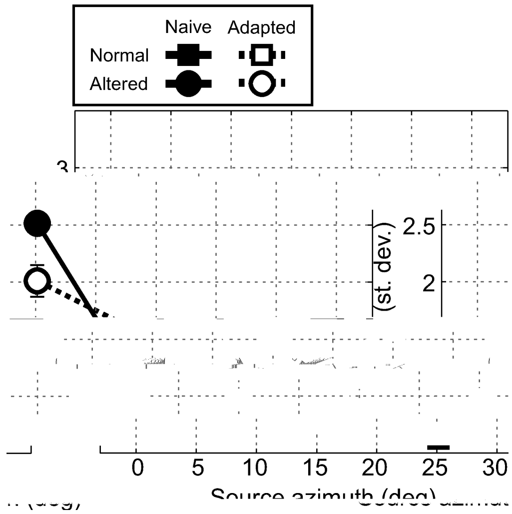

Fig. 6. Resolution (as in Figure 5) from a previous, nonlinear adaptation experiment using a comparable range of stimuli and experimental paradigm, but nonindividualized HRTFs (from [Shinn-Curringham et al. 1998a]).

may be that the subjects performing the current experiment were more skilled than those in the earlier experiment. However, it is interesting to note that individualized HRTFs were used in the current experiment, whereas a set of “generic” HRTFs was used for the subjects in the previous experiment, a fact that may have led to better spatial resolution overall in the current results.

When altered cues are first introduced, resolution increases for the center positions compared to using normal cues. This is not surprising, in that “adjacent” stimuli are actually generated using HRTFs with twice the normal angular separation. However, resolution is poor at the edges using altered cues. This result, which also arises in the results from the previous experiment, is most likely due to edge effects. In the identification task, if the source is perceived as coming from outside the range of allowed responses, the expected mean response for the location will be the edge. Such an edge effect makes it difficult to estimate d’ for positions at the edges and this effect is exacerbated when overall resolution is good, as in the current experiment.

Following training with altered cues, resolution improves in the current experiments. This result is inconsistent with results from a number of previous experiments [Shinn-Cunningham et al. 1998a]. It may be that for linear transformations, resolution with altered cues ollowingay ot7(.6(l3]TJTMk-6(al.)-26sCe1.2 tw8.

subject. As in previous experiments, resolution for "normal" cues decreased when subjects were trained to expect a larger-than-normal range of stimuli. However, unlike previous experiments, resolution using "altered" cues actually increased as subjects learned to attend to the larger-than-normal range of stimuli. Additional experiments are necessary to examine how resolution is affected by training with linear-cue transformations. In particular, subjects adapted so well to the transformation that it is difficult to estimate resolution. Results were further confounded by the fact that in the first "altered-cue" run, edge effects (inherent in the identification paradigm) made it nearly impossible to discern differences in mean response; this later effect, which is accentuated by the overall better-than-expected performance, causes the resolution to be consistently underestimated for lateral source positions in the initial altered-cue run. In order to test whether resolution using altered cues improves with training obtains generally or is an artifact of the experimental paradigm, additional experiments are planned which (1) use more source locations to improve resolution estimates and (2) employ an analog localization response rather than an identification task to get a more direct measure of the distribution of spatial percepts elicited by different stimuli during the course of the experiment.

6. DISCUSSION

Subjects can adapt to linear transformations of auditory space more completely than to more complex transformations, given relatively short training times. However, they can adapt to very complex rearrangements with sufficient training [Hofman et al. 1998]. Thus, the fact that with short-term training, subjects can only adapt to linear transformations suggests that spatial auditory plasticity occurs at many different stages in the computational pathway. In particular, short-term training may not change how spatial cues, such as interaural differences and spectral cues are computed and combined to form spatial percepts, but only how these percepts are mapped to exocentric space. Such a model implies that internal noise in the spatial percept is fixed over short time scales, a view which may explain the very good performance of subjects using individualized HRTFs compared to those using nonindividualized HRTFs. While previous results have shown that the number of gross localization errors is reduced when individualized HRTFs are used in virtual auditory displays [Wenzel et al. 1993], current results suggest that use of individualized HRTFs also decreases variability in localization judgments. This result is consistent with the idea that the computation of source location is optimized for "normal" experience; presenting unnatural combinations of cues should yield a more diffuse sound image that

SHINN-CUNNINGHAM, B. G., DURLACH, N. I., AND HELD, R. M. 1998a. Adapting to supernormal auditory localization cues. I. Bias and resolution. *J a f e Ac ca S c e f A e ca 103*, 6, 3656–3666.

SHINN-CUNNINGHAM, B. G., DURLACH, N. I., AND HELD, R. M. 1998b. Adapting to supernormal auditory localization cues. II. Constraints on adaptation of mean response. *J a f e Ac ca S c e f A e ca 103*, 6, 3667–3676.

WENZEL, E. M., ARRUDA, M., KISTLER, D. J., AND WIGHTMAN, F. L. 1993. Localization using nonindividualized head-related transfer functions. *J a f e Ac ca S c e f A e ca 94*, 111–123.

WIGHTMAN, F. L. AND KISTLER, D. J. 1989. Headphone simulation of free-field listening. I. Stimulus synthesis. *J a f e Ac ca S c e f A e ca 85*, 858–867.

ZAHORIK, P.