Object continuity enhances selective auditory attention

B[illegible], E J. O[illegible], N[illegible] K[illegible], B[illegible] G. S[illegible]-C[illegible]*

Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, MA 02215

Edited by Eric I. Knudsen, Stanford University School of Medicine, Stanford, CA, and approved June 27, 2008 (received for review April 16, 2008)

[illegible]

source segregation | auditory scene analysis | spatial hearing | streaming | auditory mixture

In everyday situations, we are confronted with multiple objects that compete for our attention. Both stimulus-driven and goal-related mechanisms mediate the between-object competition to determine what will be brought to the perceptual foreground (1, 2). In natural scenes, objects come and go and the object of interest can change from moment to moment, such as when the flow of conversation shifts from one talker to another at a party. Thus, our ability to analyze objects in everyday settings is directly affected by how switching attention between objects affects perception. Much of what we know about the effects of switching attention comes from visual experiments in which observers monitor rapid sequences of images or search for an item in a static field of objects (3, 4). Although these situations give insight into the time it takes to dis- and reengage attention from one object to the next, they do not directly explore whether there are dynamic effects of sustaining attention on one object through time.

In contrast to visual objects, the identity of an auditory object

We predicted that providing spatial information in advance during the gaps between digits in the target sequence would eliminate the cost of switching spatial attention. In the ''switching, LED leading (SL)'' condition, the LEDs were turned on at the beginning of the silent gap preceding a target digit (see *Materials and Methods*

quality) or a rapid presentation rate. Thus, parsimony favors the hypothesis that selective attention becomes increasingly more effective as object formation builds.

When the target sequence has spatial continuity and maximal voice continuity (Fig. 3 *Lower*, leftmost plot), performance for the first digit in the sequence is better than when spatial location changes between digits. This kind of effect can only be explained if the overall difficulty of a trial impacts how well the first digit of the target sequence is recalled at the conclusion of the trial, because the subject has no advance knowledge about the target location or target voice for the first digit in either the F or SS conditions. This result suggests that attentional demands are smallest when the target sequence is temporally connected, continuous in voice quality, and from a fixed location, leaving more resources for storage and recall of the sequence. This effect undoubtedly depends on overall memory demands of the task, and thus is likely to vary with the length of the target sequence as well as the listener's knowledge about when the sequence will end.

These findings shed light on why, in listening environments such as noisy parties or restaurants, it is more difficult to follow a conversation involving many people (where the relevant talker often and unexpectedly changes locations) than to focus on one talker (at one location) exclusively6) exclthus isoneone shedone tem)-54-2s and kare25).er

were spoken by the same voice (chosen randomly on each trial). The maskers were chosen from the remaining 14 voices (separately for each temporal position).

P .