Perceptual grouping in a large-scale neural network model of auditory pattern recognition

Antonio Ulloa, Fatima Husain, Barry Horwitz
Language Section, NIDCD, NIH, Bethesda MD 20892

February 8, 2002

Abstract

Perceptual grouping allows the auditory system to integrate brief, disparate sounds into cohesive perceptual units. This is useful in separating attended sounds from environmental noise. Perceptual grouping was investigated with a large-scale computational model of auditory pattern recognition in a same-different short-term memory task. The model was able to match a tonal pattern with fragmented versions of the same pattern. Perceptual grouping in the model broke down gradually as the fragmentation increased. These results are explained in the context of feedforward connections among the modules of the model. We propose that such processes occur in the central auditory system.

Introduction

Perceptual grouping allows the auditory system to integrate sounds into perceptual units, even when those sounds are separated in time, but similar or continuous along some other dimension. Perceptual grouping is particularly useful for separating attended sounds (coming from the same source) from environmental noise and from other sounds. This ability of the auditory system has also been called stream segregation (Bregman, 1990).

In the present study, a computational model of auditory pattern recognition, previously presented by Husain et al. (2001), is shown to reproduce the phenomenon of perceptual grouping of auditory stimuli. The auditory model is based on neurophysiological and neuroanatomical data from primate and human studies, including functional neuroimaging studies, and was developed to integrate data from disparate studies and to simulate PET/fMRI activity. The auditory model consists of 5 modules: MGN, similar to medial geniculate nucleus; Ai, similar to the core area of auditory cortex; Aii, corresponding to belt and parabelt areas of auditory cortex; STG, corresponding to superior temporal gyrus; and PFC, corresponding to prefrontal cortex. These five modules, which are composed of multiple excitatory-inhibitory units, are linked by feedforward and feedback connections. The MGN module is the first stage in the processing of auditory inputs. Units of Ai and Aii exhibit a tonotopic organization, and have two types of direction selectivity: up-selective units responsive to sweeps increasing in frequency and down-selective units responsive to sweeps decreasing in frequency. Additionally, Aii has units responsive to sweeps that change their direction (contour units). The STG module processes higher-level features of the stimuli and integrates them into a percept. The PFC module incorporates short-term working memory into the system. Response units in the PFC module constitute the decision center in the model. If a given number of response units exceeds a preestablished threshold, the model is said to signal a match between two stimuli. A mismatch is signaled otherwise. There are different windows of spectral (frequency) and temporal (time) integration in the first three stages of the model, with Ai having the smallest temporal window and STG the longest temporal window. This model of auditory pattern recognition, with the same parameters as those used in Husain et al. (2001), was able to match a frequency-modulated tonal pattern with a fragmented version of the same tonal pattern. This perceptual grouping ability of the model, which resembled auditory perceptual grouping phenomena, broke down gradually as the degree of fragmentation of the input stimuli increased. These results are explained in the context of feedforward connections among the different modules of the model. We propose that a similar process takes place in the central auditory system.

Methods

A same-different short-term memory task was simulated using the model described above. Stimuli presented to the model were changing tonal patterns composed of upwards and downwards frequency-modulated sweeps. The model compared a stimulus against a fragmented version of the same stimulus. The tonal patterns were introduced into the MGN module as time-varying inputs, activating only those units that were assumed sensitive to the input frequency. Fragmented tonal patterns were generated by introducing silent intervals in the activations of MGN units. A simulated experimental trial consisted of the next sequence of events: (A) Stimulus 1 (500-ms). (B) Interstimulus interval (1000 ms). (C) Stimulus 2 (500-ms). (D) Response interval (2000 ms).

Each stimulus was divided into ten 50-ms segments. Stimulus 1 was the same across all the experiments. Stimulus 2 was gradually fragmented by systematically turning off segments. Each segment of the tonal pattern used in these experiments activated a single unit in the input percept in MGN, except for the first and last segments, which activated two units each. The MGN module was composed of a total of 81 units, each unit corresponding to a specific input frequency. The location of the activated units for the stimuli in the present experiments was chosen arbitrarily, and the results here presented were independent of this choice.

Five experiments were run in which the following conditions gradually fragmented stimulus 2: (1) No segments off (intact stimulus). (2) Segments 2 and 9 off. (3) Segments 2, 4, 7, and 9 off. (4) Segments 2, 3, 4, 7, 8, and 9 off. (5) Segments 2, 3, 4, 7, 8, and 9 off; segments 1 and 10 decreased to 1 unit.

The number of active units was determined by a computer program that counted the number of units exceeding a threshold level of 0.6 (1 is the maximum activation level) in the response module. A match was signaled if there was above-threshold activation in at least 5 units. A mismatch was signaled otherwise.

Results

The number of active units in the response module, for each of the conditions enumerated above, was: (1) 10, (2) 9, (3) 7, (4) 6, (5) 4. Experiments 1-4 signaled a match. Experiment 5 signaled a mismatch.

Discussion

The auditory model, with the same parameters as in previous simulations, was able to match the fragmented stimulus with the normal stimulus by performing perceptual grouping. The model was tested further by increasing the degree of fragmentation of the stimulus, until the model was not able to recognize them as a match. The perceptual grouping ability of the model broke down gradually as the degree of fragmentation increased. The finding that the model is able to match a stimulus with a fragmented version of the same stimulus can be explained in the context of the feedforward connections among the different modules. These feedforward connections in the first three stages of the model (Ai, Aii, and STG) spread excitation activity –topologically– to more than one specific unit, thereby filling out the gaps in adjacent units of the fragmented stimulus. We propose that a similar process takes place in the central auditory system.

References

Bregman, A. (1990). *Auditory scene analysis: The perceptual organization of sound.* The MIT Press, Cambridge, MA.

Husain, F., Tagamets, M.-A., Braun, A., and Horwitz, B. (2001). Large-scale computational model for simulating PET/fMRI studies of auditory pattern recognition. *NeuroImage*, 13(S1307).