

Supplementary Materials for  
**Speaker-independent auditory attention decoding without access to  
clean speech sources**

Cong Han, James O'Sullivan, Yi Luo, Jose Herrero, Ashesh D. Mehta, Nima Mesgarani\*

\*Corresponding author. Email: [nima@ee.columbia.edu](mailto:nima@ee.columbia.edu)

Published 15 May 2019, *Sci. Adv.* **5**, eaav6134 (2019)  
DOI: 10.1126/sciadv.aav6134

**The PDF file includes:**

Fig. S1. Electrode coverage and speech responsiveness for each subject.

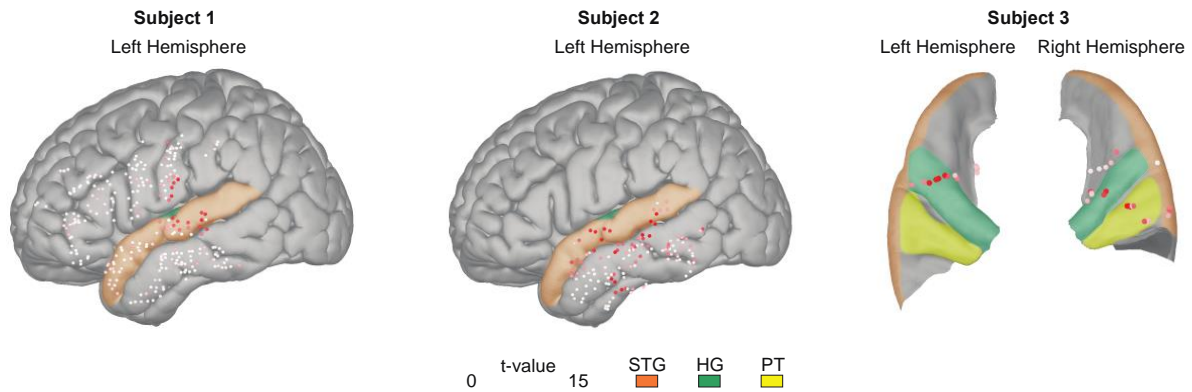
Fig. S2. The change in the update parameter of attractors (parameter  $q$  in methods) when the speakers in the mixture switch.

Legend for movie S1

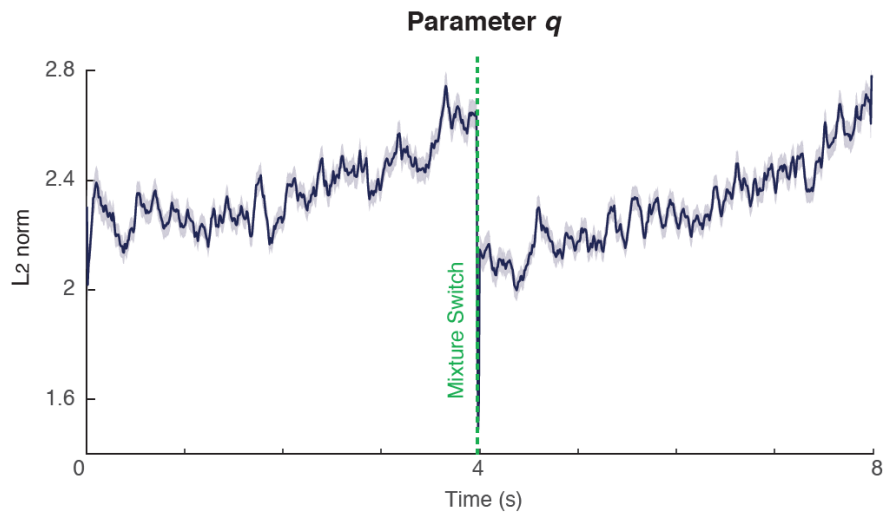
**Other Supplementary Material for this manuscript includes the following:**

(available at [advances.sciencemag.org/cgi/content/full/5/5/eaav6134/DC1](https://advances.sciencemag.org/cgi/content/full/5/5/eaav6134/DC1))

Movie S1 (.mov format). The full demo of the proposed ODAN-AAD system.



**Fig. S1. Electrode coverage and speech responsiveness for each subject.** Subjects 1 and 2 were implanted with high-density subdural electrode arrays over their left (language dominant) temporal lobe with coverage over the superior temporal gyrus (STG; orange). Subject 3 partook in stereotactic EEG (sEEG) in which they were implanted bilaterally with depth electrodes. These differences in implantation resulted in varying coverage of the STG, Heschl's gyrus (HG; green) and planum temporale (PT; yellow) in the left and right auditory cortices. The t-value resulting from t-test between speech versus silence is plotted on a red color scale.



**Fig. S2. The change in the update parameter of attractors (parameter  $q$  in methods) when the speakers in the mixture switch.** The parameter automatically decreases the contribution of the past attractor estimates after a change in the speakers.

### Supplementary Video 1

**Movie S1. The full demo of the proposed ODAN-AAD system.** The subject is listening to a mixture of two speakers. The subject first attends to the male speaker, and then switches to the female speaker after 25 seconds. The system is turned on after 10 seconds, and quickly detects and amplifies the attended speaker relative to the background. The target speaker is amplified by 20dB to show the effect more clearly. The plot shows the likelihood of target speaker over time.