Neural correlates of perceptual constancy in Auditory Cortex

Stephen M Town, Katherine C Wood & Jennifer K Bizley

Ear Institute, University College London, 332 Gray's Inn Road, London, WC1X 8EE, UK

Corresponding author: Stephen Town (s.town@ucl.ac.uk)

Supplemental Information

Fig. S1 Electrode positions in auditory cortex

(A) Electrode positions in two representative animals in auditory cortex. Green and black indicate electrodes within or outside (excluded from analysis) Auditory Cortex respectively. Labels show suprasylvian sulcus (sss) and pseudosylvian sulcus (pss). **(B)** Cresyl-violet stained electrode tracks.

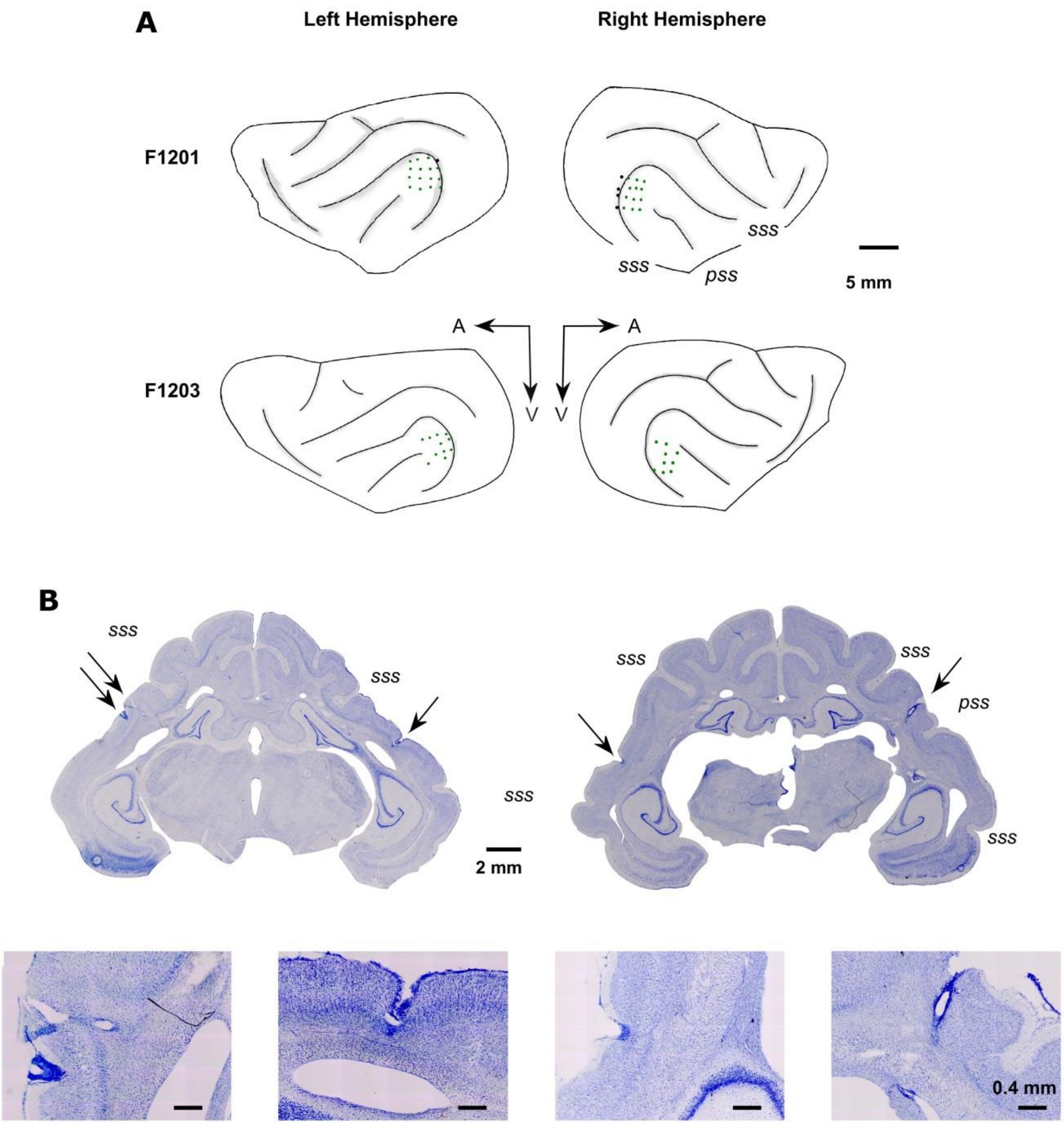
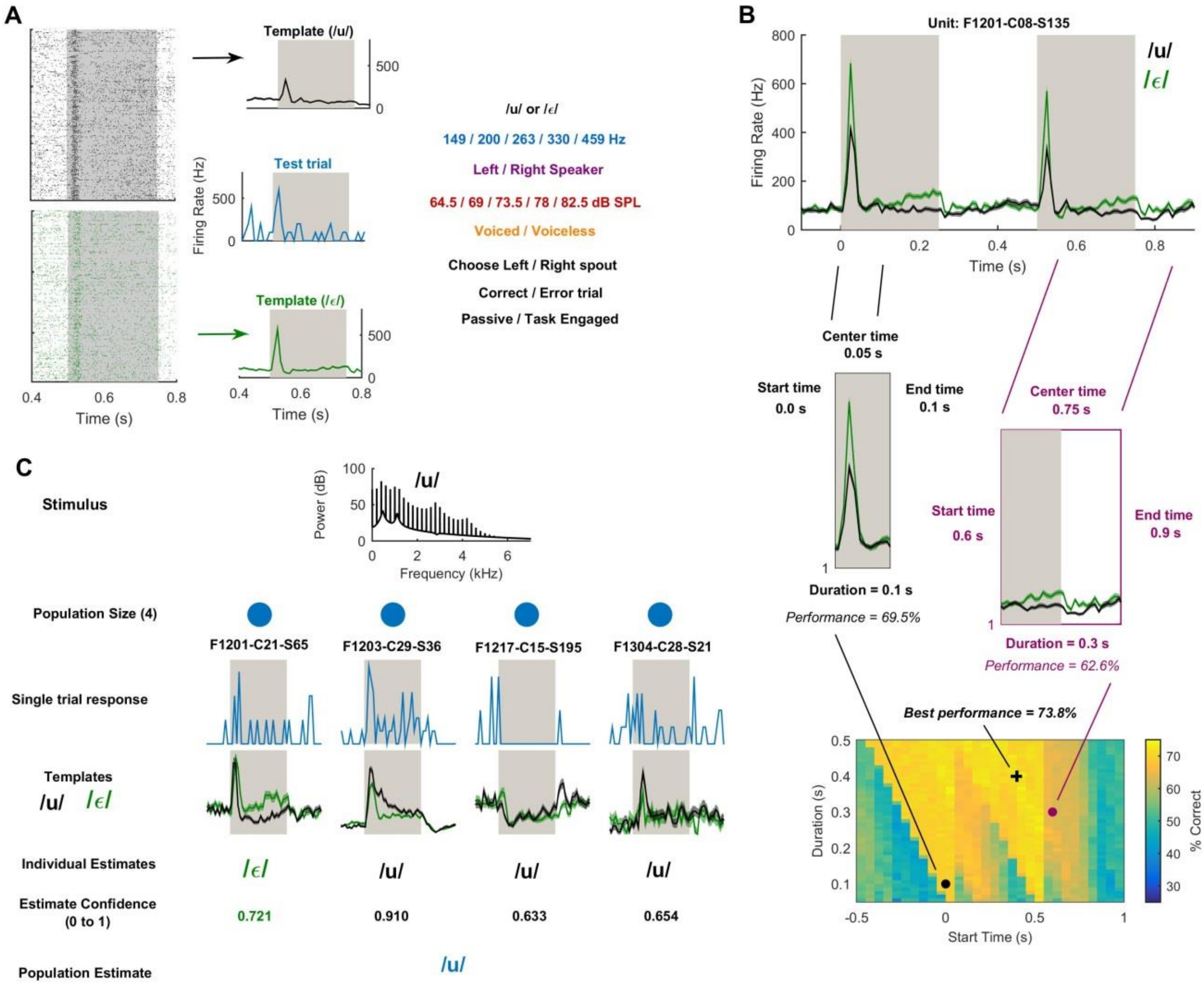


Fig. S2 Decoder structure

(A) Decoding trial parameters from single trial neural responses. For each data set recorded from a given neuron (e.g. in which vowel identity varied across F0) we used a leave-one-out cross validation method in which templates were calculated as the mean response to each stimulus class (e.g. vowel) on all but one test trial of the data set. Mean responses were averaged across trials from spike times within a decoding window binned at 10 ms intervals. For the test trial, the decoder estimate of stimulus class was assigned as the template class with the smallest Euclidean distance. Every trial in the dataset was decoded as a test trial with templates recalculated from all other trials. **(B)** To accommodate potential variation in timing of information content, we varied the temporal parameters (start time and duration) that defined the decoding window. Start time was varied from -0.5 to 1 s after stimulus onset in 50 ms intervals; duration was varied between 10 and 500 ms in 10 ms intervals. For each combination of start time and duration, we calculated decoding performance across trials and mapped temporal parameter space using a simple grid search. While this search protocol may not find the true optimal parameters for best decoding performance, it nonetheless enabled us to improve decoding performance and estimate those times in the trial at which information about a given feature was most strongly represented. **(C)** Population decoding in which single trial estimates of stimulus features (e.g. vowel identity) in optimized decoding time windows were combined across units. For the population estimate, individual units were weighted by a confidence metric defined as the distance from test trial to estimated template expressed as a proportion of the sum of distances between test trial and all templates.



A
Template (/u/)
Test trial
Template (/ϵ/)
Firing Rate (Hz)
Time (s)
/u/ or /ϵ/
149 / 200 / 263 / 330 / 459 Hz
Left / Right Speaker
64.5 / 69 / 73.5 / 78 / 82.5 dB SPL
Voiced / Voiceless
Choose Left / Right spout
Correct / Error trial
Passive / Task Engaged
B
Unit: F1201-C08-S135
/u/
/ϵ/
Firing Rate (Hz)
Time (s)
Center time
0.05 s
Start time
0.0 s
End time
0.1 s
Duration = 0.1 s
Performance = 69.5%
Center time
0.75 s
Start time
0.6 s
End time
0.9 s
Duration = 0.3 s
Performance = 62.6%
Best performance = 73.8%
Duration (s)
Start Time (s)
% Correct
C
Stimulus
/u/
Power (dB)
Frequency (kHz)
Population Size (4)
F1201-C21-S65
F1203-C29-S36
F1217-C15-S195
F1304-C28-S21
Single trial response
Templates
/u/ /ϵ/
Individual Estimates
/ϵ/
/u/
/u/
/u/
Estimate Confidence
(0 to 1)
0.721
0.910
0.633
0.654
Population Estimate
/u/

Fig. S3 Decoder performance vs. Number of classes

(A-B) Decoding vowel identity and task-orthogonal values for stimuli at the most extreme values of fundamental frequency (**A**) and sound level (**B**) over which units were simultaneously tested. (The highest sound level tested was 82.5 dB SPL but the lowest level for such tests was 64.5 dB SPL). **(C-D)** Percentage of units informative about vowel identity and / or task-orthogonal values for sounds varied over two F0s separated by ΔF0 **(C)** or two sound levels separated by ΔLevel **(D)**. Dots indicate individual comparisons; lines show regressions between feature separation and proportion of informative units. The distribution of unit classes was consistent across changes in ΔF0 whereas increasing separation of sound levels led to an increase in the proportion of units informative about only sound level (red) and both sound level and vowel identity (black). As discrimination of sound level improved, more units moved from being informative about vowel identity only (green) to being informative about level and vowel identity. Thus the number of units informative about vowel identity *alone* decreased with ΔLevel, but the number of vowel informative units remained constant when information about sound level as not considered (green + black lines).

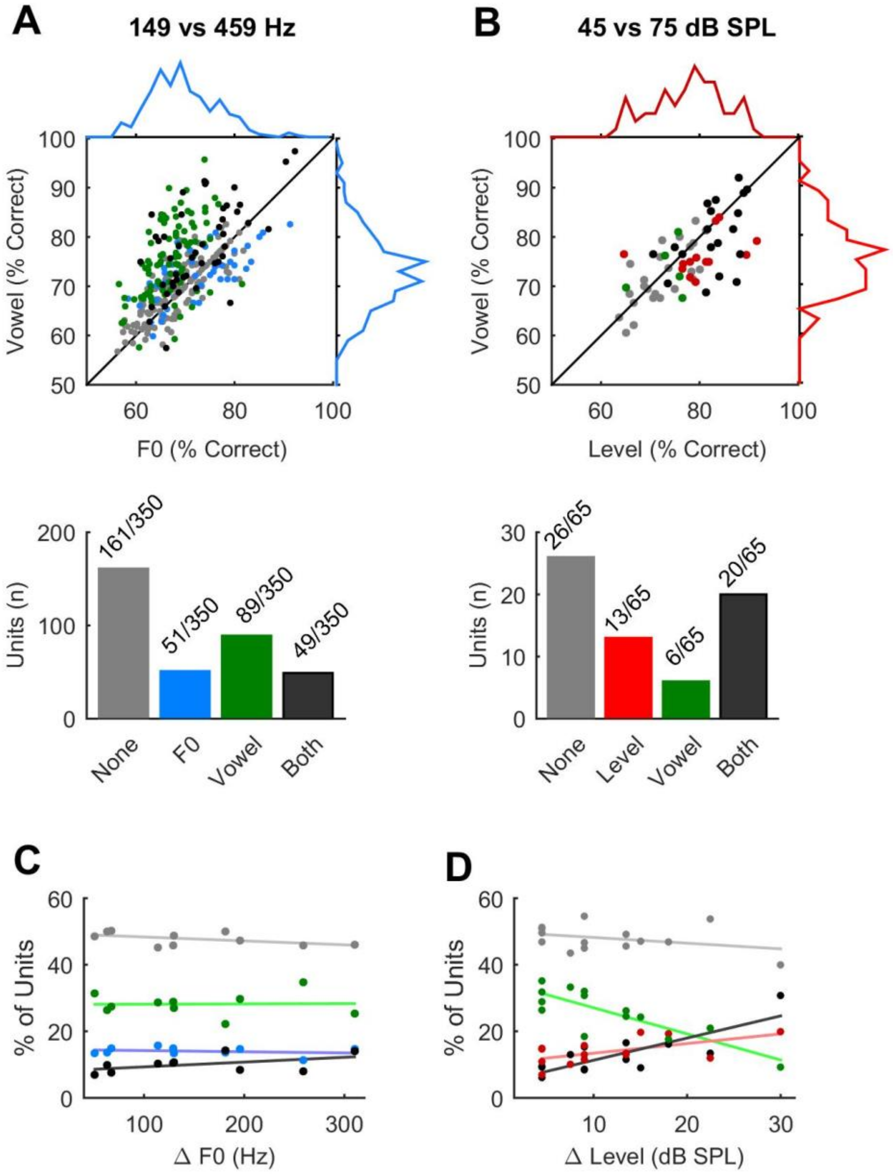

A
149 vs 459 Hz
B
45 vs 75 dB SPL
Vowel (% Correct)
F0 (% Correct)
Level (% Correct)
Units (n)
161/350
51/350
89/350
49/350
26/65
13/65
6/65
20/65
None
F0
Vowel
Both
Level
C
D
% of Units
Δ F0 (Hz)
Δ Level (dB SPL)

Fig. S4 Vowel decoding across individual sound levels – alternate level range

(A) Decoding performance and unit classification shown for sounds roved in level between 64 to 82 dB SPL and 45 to 75 dB SPL. Decoding of vowel identity did not differ between units tested with each sound range (t_{192} = -1.50, p = 0.136) and the proportion of units informative about vowel identity was similar for sounds separated by 7.5 dB (25/65, 38.5%) and 4.5 dB (53 / 129, 41.1%). Decoding of sound level was significantly better for sounds separated by 7.5 than 4.5 dB (two-sample t-test, t_{192} = -4.90, p = 2.03 x 10^{-6}) indicating that neural responses were more discriminable when separation of sound levels increased. This was also reflected in a greater proportion of units informative about sound level (permutation test, $p < 0.05$) with sound level intervals of 7.5 dB (23/65, 35.4%) than 4.5 dB (25 / 129, 19.4%). **(B)** Population decoding of vowel identity and sound level shown separately for units tested with sounds with different sound level ranges. **(C)** Pair-wise comparison of decoder performance for decoding vowel identity at specific sound levels that were ranged from 45 to 75 dB SPL. Data for sounds ranged from 64 to 82 dB are shown in figure 3C of the main text. Performance was most closely correlated with similar sound levels and correlation coefficients decreased as a function of sound level difference (linear regression, F_8 = 12.4, p = 0.00783). **(D)** Temporal multiplexing of information about sound level and vowel identity shown separately for units tested with different sound ranges. Vowel identity information precedes sound level in all informative units; however while in the main text differences between vowel and sound level timing are significant, no significant differences are detected when the population of units is subdivided.

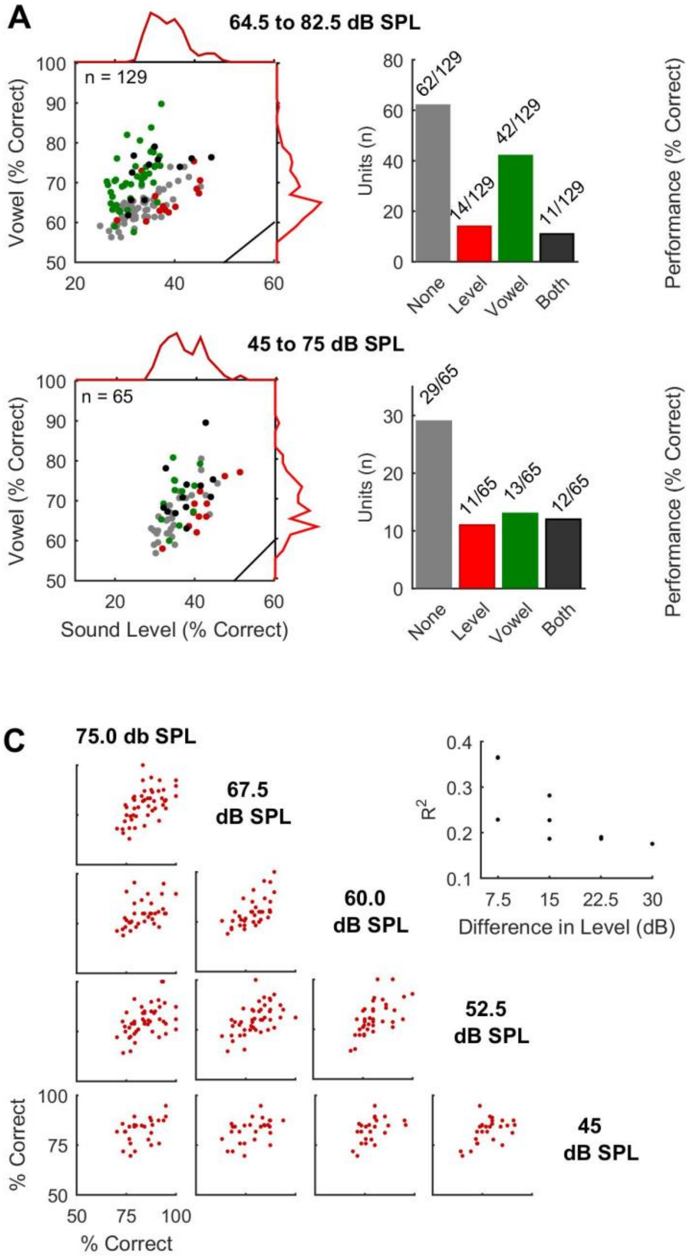

A
64.5 to 82.5 dB SPL
n = 129
Vowel (% Correct)
Units (n)
62/129
14/129
42/129
11/129
None
Level
Vowel
Both
45 to 75 dB SPL
n = 65
Sound Level (% Correct)
29/65
11/65
13/65
12/65
B
Vowel Identity
Level
Performance (% Correct)
Population size
C
75.0 db SPL
67.5 dB SPL
60.0 dB SPL
52.5 dB SPL
45 dB SPL
% Correct
Difference in Level (dB)
D
Uninformative
Single feature
Dual feature
% of Units
p = 0.210
p = 0.715
p = 0.487
p = 0.379
p = 0.092
p = 0.278
Center Time (s)

Fig. S5 Temporal profiles for the onset (start time) of best decoding windows

(A-D) Cumulative distributions showing start time for best performance when decoding vowel identity or task-orthogonal variables (**A:** F0, **B**: location, **C**: level and **D**: voicing). Units shown separately by classification as informative about vowel identity and task-orthogonal values (Dual feature units), either vowel identity or task task-orthogonal (Single feature units) or neither feature (uninformative). **(E)** CDFs for decoding vowel identity across each task-orthogonal variable. **(F)** CDFs for decoding task-orthogonal values across vowels.