SUPPLEMENTARY RESULTS

Behavioral Measures

Independent samples *t*-tests showed musicians to have significantly higher tone identification [t (18) = 3.664, $P < 0.005$] and tone discrimination [t (18) = 3.224, $P < 0.005$] scores than nonmusicians. The mean (and standard error) tone identification scores for musicians and nonmusicians are 0.9090 (0.0223) and 0.7450 (0.0388), respectively. The mean (and standard error) tone discrimination scores for musicians and nonmusicians are 0.9000 (0.0132) and 0.7480 (0.0450), respectively.

Pitch-Tracking Measures

Because Pearson's correlation coefficients do not comprise a normal distribution, *stimulus-to-response correlation* and *autocorrelation* measures were converted to Fisher's z' scores before subsequent parametric statistical analyses.

In terms of *stimulus-to-response correlation*, a 3 (tone) x 2 (group) repeated measures ANOVA revealed main effects of tone [F (2, 36) = 17.985, $P < 0.001$] and group [F (1, 18) = 8.039, $P < 0.015$] but no significant interaction. Thus, musicians showed more precise pitch-tracking compared to nonmusicians. To examine in detail any possible tone differences between groups, we performed one independent samples *post hoc t*-tests and found no pitch-tracking difference in Tone 1 but marginally significant group differences in Tone 2 [t (18) = 2.237, $P = 0.019$], and significant group difference in Tone 3 [t (18) = 2.355, $P = 0.015$] (per Bonferroni procedures, $P < 0.016$ is required for establishing statistical significance for the three tests performed: 0.05/3 = 0.01667),

suggesting that the main effect of group was largely driven by Tone 3 and less so by Tone 1 and Tone 2.

Likewise, *autocorrelation* values were entered into a 3 x 2 repeated measures ANOVA, which showed a main effect of tone [F (2, 36) = 14.135, $P < 0.001$], but no main effect of group [F (1, 18) = 1.819, $P = 0.194$]). A marginally significant group x tone interaction was found [F (2, 36) = 2.741, $P = 0.078$]. Based on the visual inspection of the autocorrelation plots showing musicians to have more robust responses on Tone 3 but not the other tones, a one independent samples t-test was performed on Tone 3, which confirmed musicians' higher *autocorrelation* values [t (18) = 2.059, $P = 0.027$]. Overall, these pitch-tracking measures showed musicians to have more robust neural phase-locking and more faithful pitch-tracking. This group difference was particularly pronounced in Tone 3, the most acoustically complex tone.

In addition to the above pitch-tracking measures, we also compared musicians' and nonmusicians' f_0 *amplitude* for each tone to examine whether the pitch-tracking differences discussed above could also be attributed to musicians' stronger FFR. A repeated measures ANOVA revealed a main effect of group (stronger in musicians) [F (1, 18) = 4.385, $P = 0.051$], a main effect of tone [F (2, 36) = 3.315, $P < 0.05$] but no significant interaction. *RMS amplitude* of the FFR waveform was likewise compared. We found a main effect of group (stronger in musicians) [F (1, 18) = 4.329, $P = 0.052$], a main effect of tone [F (2, 36) = 6.693, $P < 0.005$], but no significant interaction. Moreover, to examine whether f_0-*FFR proportion* differed between the two subject groups, a series of Mann-Whitney U-tests (on each tone) were performed and no

significant group difference was found. Non-parametric tests were used because of the non-normal distribution of the data derived from this measure.

Correlations between Pitch-Tracking Measures and Music Background

To investigate the relationship between pitch-tracking measures and musical experience, a series of Pearson's correlations were calculated. We calculated correlations between each pitch-tracking measure and years of musical training, as well as age at which musical training began (age onset). Four of the subjects had no formal musical training, and thus were not included in the correlation analyses involving age onset. Given that Tone 3 is the only tone shown to differ consistently between musicians and nonmusicians, only measures derived from this tone were considered. The correlations between years of musical training and *stimulus-to-response correlation* for Tone 3 and *autocorrelation* for Tone 3 are 0.456 ($P = 0.022$) and 0.549 ($P = 0.006$), respectively ($P < 0.025$ is required to establish statistical significance for the two tests performed). The correlations between age onset and stimulus-to-response correlation for Tone 3 and autocorrelation for Tone 3 are –0.502 ($P = 0.024$) and –0.332 ($P = 0.105$), respectively.