Figure captions

Figure 1 (left panel): Averaged tap-tone asynchronies for tone sequences containing negative shifts (tone presented 90 ms earlier than expected: filled circle) and positive shifts (tone presented 90 ms later than expected: unfilled circle). On the x-axis, *T0* denotes the stimulus where the shift occurred. Four regular tone sequences before (*T-4* to *T-1*) and after (*T+1* to *T+4*) the shift are shown. **Figure 1 (right panel)**: The identical data were transformed to show normalized asynchronies following a shift (*T0*), to compare the error correction performance between negative and positive shift conditions. On the y-axis, '0' indicates the baseline negative mean asynchrony (average of *T-4* to *T-1*), and '1' on the y-axis shows the maximum deviance from the baseline owing to the shift. Positive shifts were corrected faster with a degree of over-correction (unfilled circle), compared with negative shifts (filled circle), [$p < .05$]. Error bars represent standard error of mean.

Figure 2: Grand averaged stimulus-locked ERPs to all 4 conditions from FCz for illustration purpose only. ERPs were time-locked to *T-2* (at 0 ms). These macro-epochs contain preceding tones (*T-2* & *T-1*), a tone subject to a ±90 ms time-shift (*T0*), and 4 subsequent tones (*T+1* to *T+4*). Condition labels indicate the shift direction of *T0* (-ve shift: 90 ms earlier than expected or +ve shift: 90 ms later than expected).

Figure 3 (upper panel): Grand averaged ERPs from FCz showing stimulus-locked epochs to the shift position *T0* (at 0 ms on the x-axis) for listening and tapping conditions of both shift directions. A significant 2-way interaction between Condition (listening vs. tapping) and ShiftDirection (negative vs. positive) was identified in 2 time windows (shaded boxes: N1 around 100 ms [$F(1,14) = 19.77$, $p < .001$] and N2 around 300 ms [$F(1,14) = 15.06$, $p < .001$]). **Figure 3 (lower panel)**: Topographic maps for each condition for each time window, and their corresponding significance maps were shown (at 119 ms and 316 ms). Note that warmer colors represent positivity.

Figure 4: Grand averaged ERPs from FCz showing stimulus-locked epochs, time-locked to *T-1* or *T0* (0 ms on the x-axis) for tapping negative and positive conditions. ERPs were relative to the baseline period from -50 to 0 ms. A significant 2-way interaction Position (*T-1* vs. *T0*) and ShiftDirection (negative vs. positive) was identified in 2 time windows (shaded boxes: N1 around 100 ms [$F(1,14) = 31.55$, $p < .001$] and N2 around 300 ms [$F(1,14) = 25.13$, $p < .001$]).

Figure 5 (upper panel): Grand averaged ERPs from FCz, showing response-locked epochs, time-locked to the tap-onset for *T-1* or *T0* stimulus (at 0 ms on the x-axis) for tapping negative and positive conditions. ERPs were relative to the baseline period from -50 to 0 ms. No significant 2-way interaction was identified between Position (*T-1* vs. *T0*) and ShiftDirection (negative vs. positive). The shaded box (356-408 ms) indicates the significant window of ShiftDirection main effect (tapping negative condition > tapping positive condition). It was most significant at 374 ms [$F(1,14) = 6.22$, $p < .05$]. **Figure 5 (lower panel)**: Topographic maps showing activity at 374 ms. It compares tapping negative and tapping positive conditions at *T0* only. Note that warmer colors represent positivity.

Figure 6 (left panel): The peak amplitude and latency of CNV-like negativity for each participant. In the tapping negative condition, there was a significant positive correlation between the CNV-like negativity peak latency and the normalized error correction performance at *T+1* (i.e., the earlier the peak, the better the error correction performance with the negative shifts) [$r(15) = .569$, $p = .027$]. **Figure 6 (right panel)**: In the tapping positive condition, there was a trend level of negative correlation between the CNV-like negativity peak latency and the normalized error correction performance at *T+1* (i.e., the later the peak, the better the error correction performance with the positive shifts) [$r(15) = -.439$, $p = .10$].

Figure 1

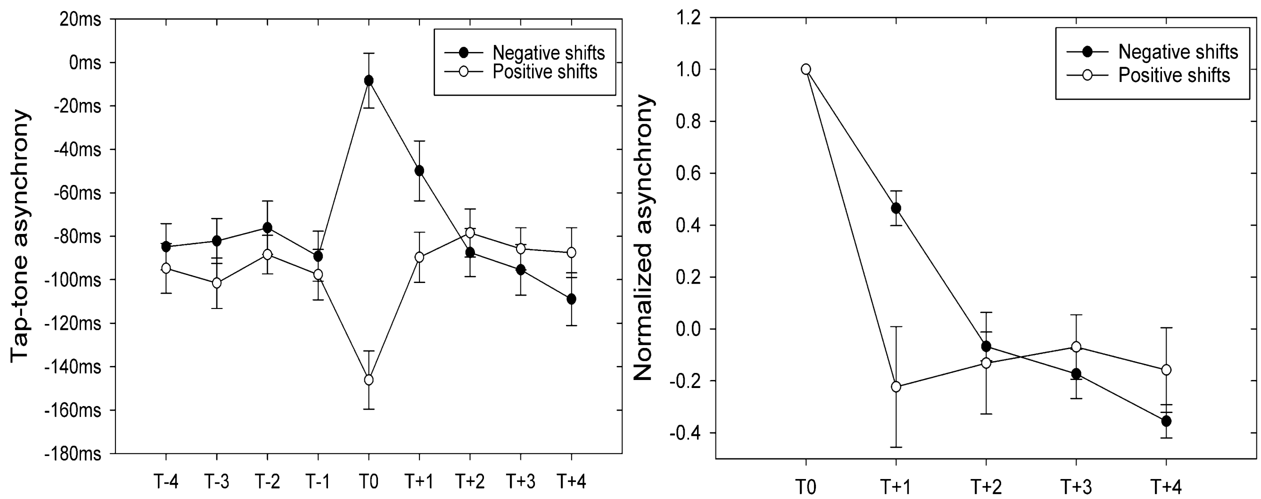
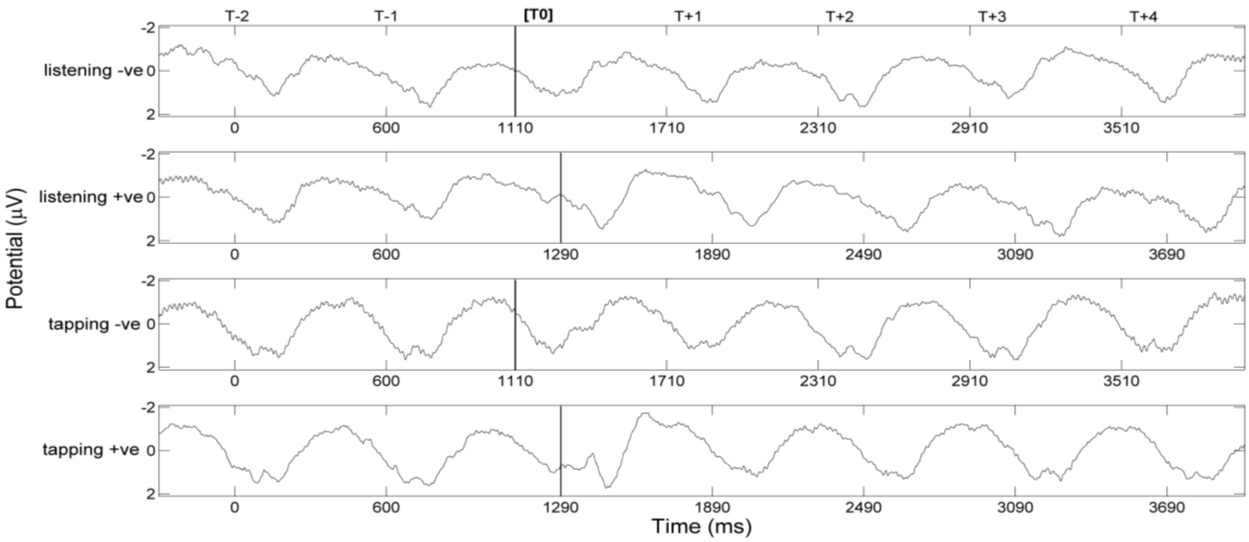
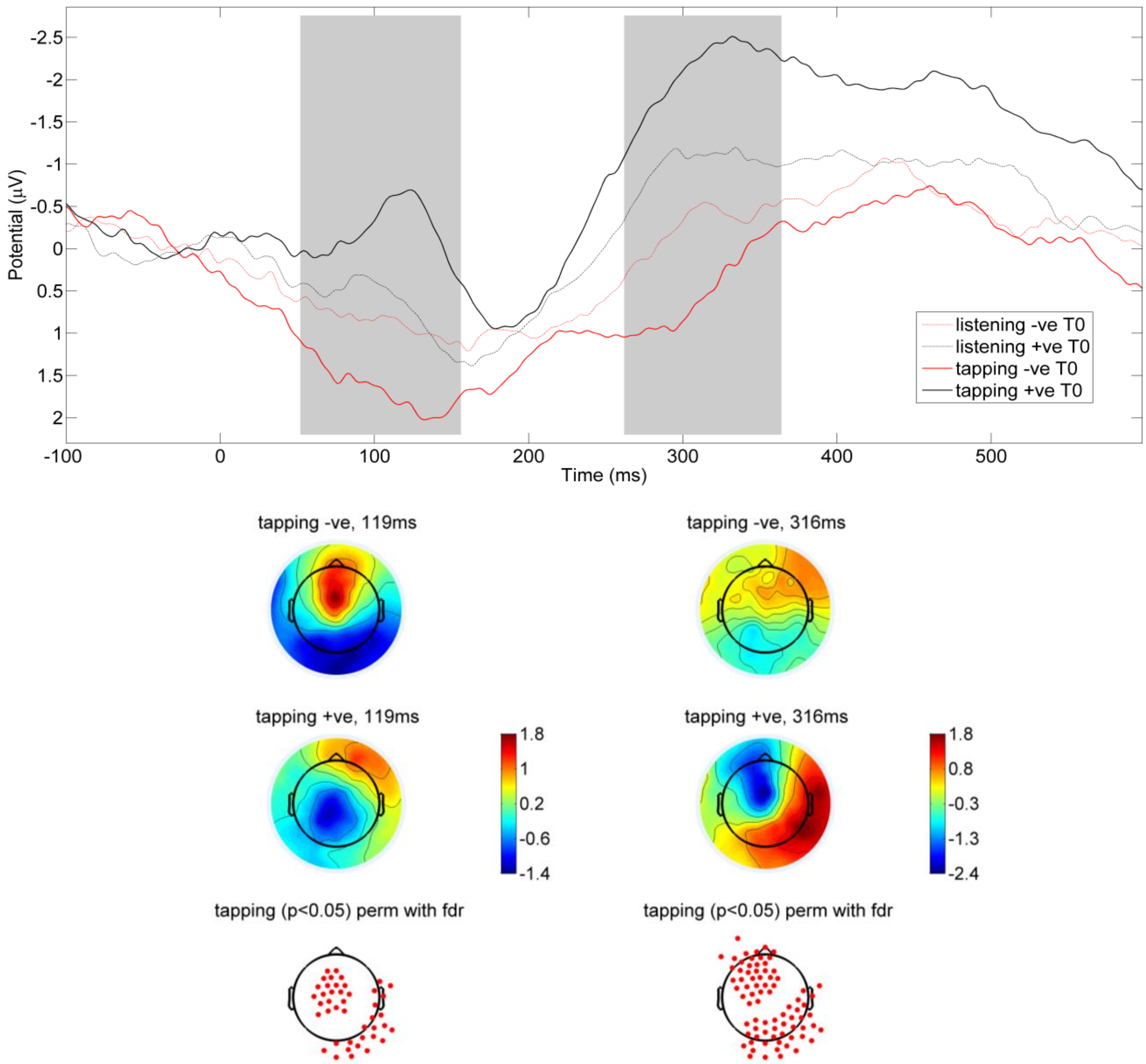

Tap-tone asynchrony
20ms
0ms
-20ms
-40ms
-60ms
-80ms
-100ms
-120ms
-140ms
-160ms
-180ms
T-4
T-3
T-2
T-1
T0
T+1
T+2
T+3
T+4
Negative shifts
Positive shifts
Normalized asynchrony
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
T0
T+1
T+2
T+3
T+4
Negative shifts
Positive shifts

Figure 2



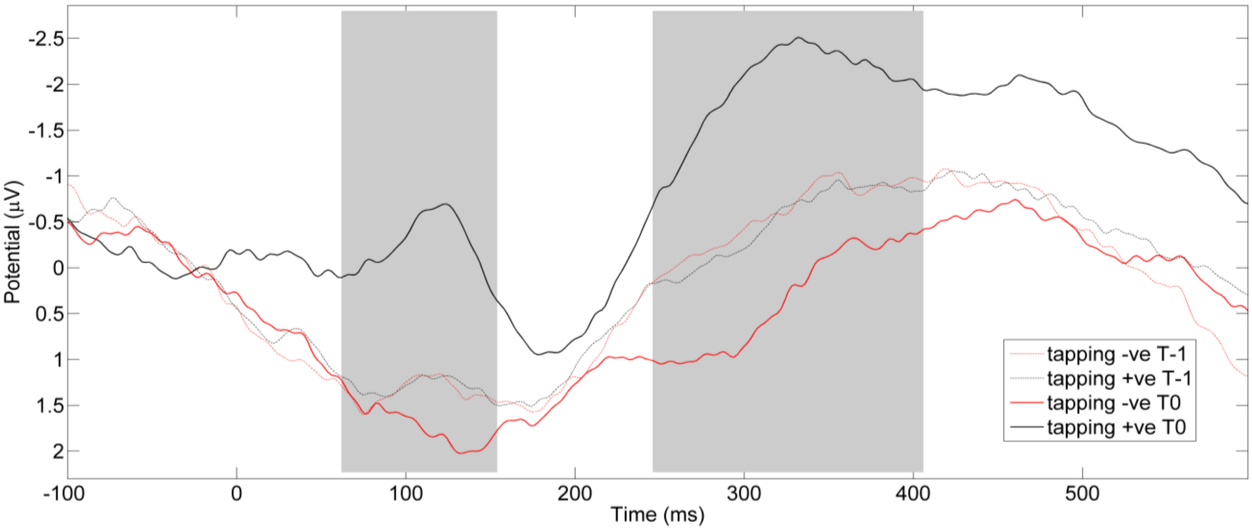
T-2
T-1
[T0]
T+1
T+2
T+3
T+4
-2
listening -ve 0
2
0
600
1110
1710
2310
2910
3510
-2
listening +ve 0
2
0
600
1290
1890
2490
3090
3690
-2
tapping -ve 0
2
0
600
1110
1710
2310
2910
3510
-2
tapping +ve 0
2
0
600
1290
1890
2490
3090
3690
Potential (μV)
Time (ms)

Figure 3



-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
Potential (μV)
-100
0
100
200
300
400
500
Time (ms)
listening -ve T0
listening +ve T0
tapping -ve T0
tapping +ve T0
tapping -ve, 119ms
tapping -ve, 316ms
tapping +ve, 119ms
tapping +ve, 316ms
1.8
1
0.2
-0.6
-1.4
1.8
0.8
-0.3
-1.3
-2.4
tapping (p<0.05) perm with fdr
tapping (p<0.05) perm with fdr

Figure 4



-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
Potential (µV)
-100
0
100
200
300
400
500
Time (ms)
tapping -ve T-1
tapping +ve T-1
tapping -ve T0
tapping +ve T0

Figure 5

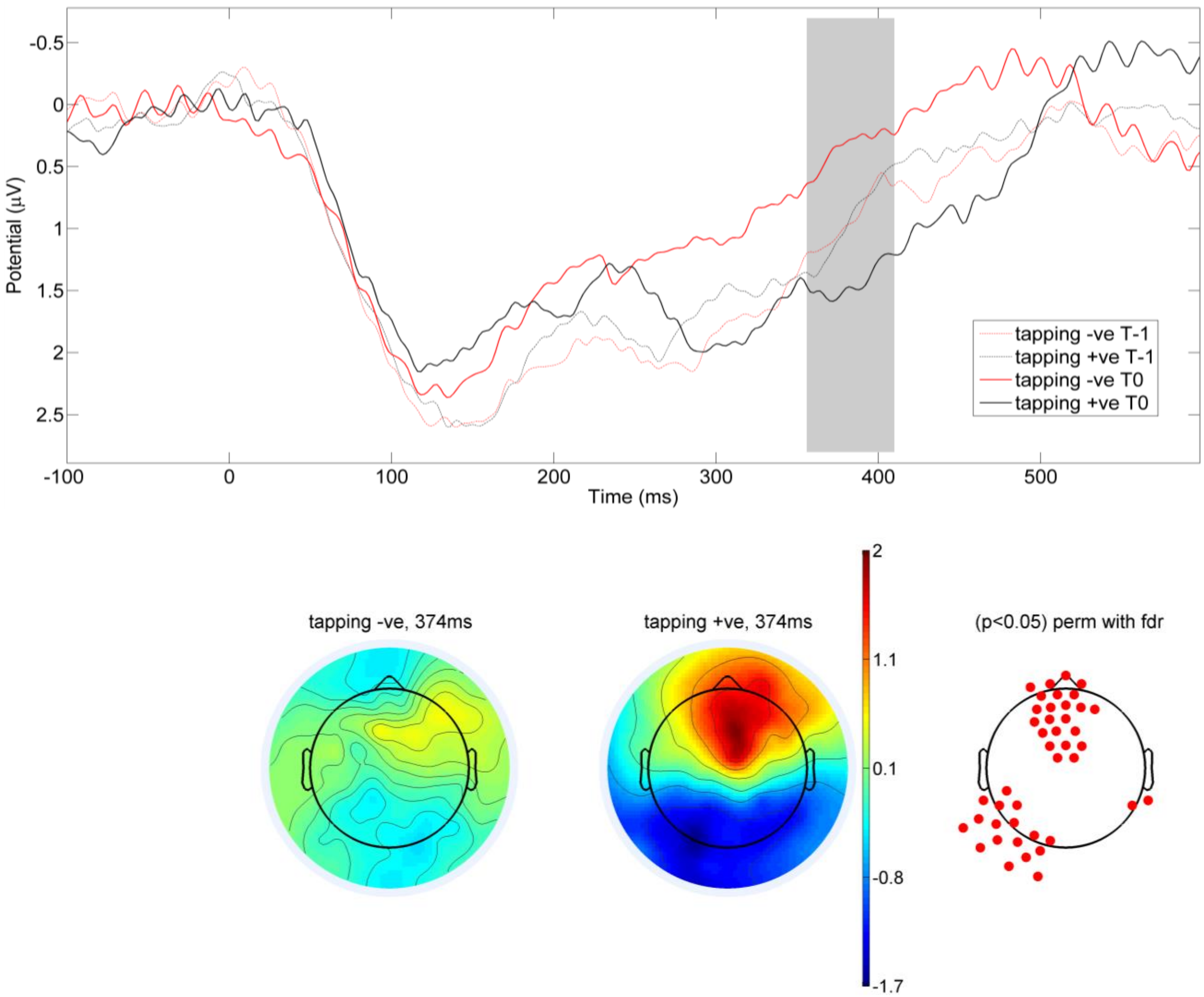
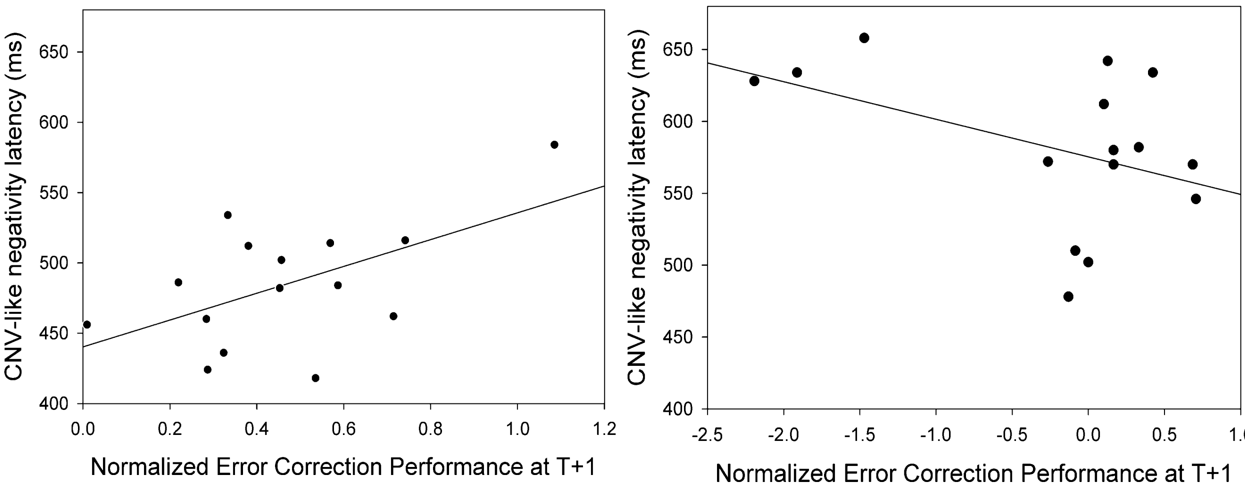

Potential (μV)
Time (ms)
tapping -ve T-1
tapping +ve T-1
tapping -ve T0
tapping +ve T0
tapping -ve, 374ms
tapping +ve, 374ms
(p<0.05) perm with fdr

Figure 6


CNV-like negativity latency (ms)
400
450
500
550
600
650
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Normalized Error Correction Performance at T+1
CNV-like negativity latency (ms)
400
450
500
550
600
650
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
Normalized Error Correction Performance at T+1