For the time window statistics, the peak for each condition's MMN was identified as the local minimum of the grand-averaged waveform at Fz within the 100-300 ms interval after the sound onset. The 51-ms time window surrounding that peak (25 ms before and after) was selected for analysis, and the mean amplitude at Fz within that time window was calculated for each participant. Note that this means different time windows were selected for each condition; the labels of the topographic maps in the EEG figures in the main paper indicate the time window used for that condition.

The mean averages were then submitted to a CONTRAST (3 within-participant levels: T3~T2, T3~T4, T2~T4; or, for Experiment 3, 6 within-participant levels) × DIRECTION (2 within-participant dummy levels) × GROUP (2 between-participant levels: Chinese speakers, non-Chinese speakers) mixed ANOVA. The Greenhouse-Geisser correction was used for effects which involved the factor CONTRAST and thus had more than one degree of freedom in the numerator.

Below we only discuss effects of CONTRAST×DIRECTION or CONTRAST×DIRECTION×GROUP, as any lower-order effects are not of theoretical interest. A main effect of CONTRAST would only indicate that some contrasts yield larger MMNs overall than others, but would not reveal an asymmetry, and thus is not if interest to the present research question. A main effect of DIRECTION would be meaningless because the levels we assigned to this factor were dummy levels with no inherent meaning (for example, $yi^3 \rightarrow yi^2$ was assigned "A" and $yi^2 \rightarrow yi^3$ "B"; if there were somehow a main effect of DIRECTION such that all the "B" conditions had greater MMNs than all the "A" conditions, this effect would not happen if we arbitrarily reversed the coding in one condition).

Experiment 1. The CONTRAST×DIRECTION interaction was significant ($F(2, 60) = 14.26$, $p < .001$), indicating that the presence or magnitude of MMN asymmetries differed as a function of contrast being tested. Significant asymmetries were present in all three contrasts (T3~T2: $t(31) = -6.31$, 95% CI = -2.83…-1.44 $p < .001$; T2~T4: $t(31) = -3.60$, 95% CI = -1.63…-0.45, $p = .007$; T3~T4: $t(31) = -8.47$, 95% CI = -4.02…-2.46, $p < .001$). However, the magnitude of the asymmetry was significantly larger for T3~T4 than for T2~T4 ($t(31) = 5.23$, 95% CI = 1.34…3.06, $p < .001$), marginally larger for T3~T2 than for T2~T4 ($t(31) = -2.73$, 95% CI = -1.92…-0.28, $p = .062$), and not significantly different between T3~T2 and T3~T4 ($t(31) = 2.46$, 95% CI = 0.19…2.02, $p = .117$). (Note, however, that these significance levels are conservatively Bonferroni-corrected; the raw p-values are 1/6 those reported here.) In short, contrasts involving T3 elicited stronger asymmetries than contrasts not involving T3.

The GROUP×CONTRAST×DIRECTION interaction was also marginal ($F(2, 60) = 2.79$, $p = .072$), suggesting a trend towards differences between the two groups. In short, for Chinese-speaking listeners all three contrasts elicited significant asymmetries (ps < .008), whereas for non-Chinese-speaking listeners the T3~T2 and T3~T4 contrasts did (ps < .001) but the T2~T4 contrast did not ($p > .99$). In short, while the asymmetries involving T3 were cross-linguistically robust, the asymmetry not involving T3 was driven by the Chinese-speaking participants.

Experiment 2. The CONTRAST×DIRECTION interaction was marginal ($F(2, 60) = 2.58$, $p = .086$), indicating that the presence or magnitude of MMN asymmetries differed as a function of contrast being tested. Significant asymmetries were present for the T3~T2 contrast ($t(31) = -3.26$, 95% CI = -2.02…-0.46, $p = .016$) and the T3~T4 contrast ($t(31) = -5.23$, 95% CI = -2.46…-1.08, $p < .001$), but not for the T2~T4 contrast ($t(31) = -1.11$, 95% CI = -1.50…0.45, $p = .832$).

The GROUP×CONTRAST×DIRECTION interaction was not significant ($F(2, 60) = 0.56$, $p = .567$).

Experiment 3. No effects reached significance in the window analysis.